UNIVERSIDADE FEDERAL DE JUIZ DE FORA
FACULDADE DE DIREITO
PROGRAMA DE PÓS-GRADUAÇÃO *STRICTO SENSU* EM DIREITO

Eduardo Khoury Alves

Privacidade à exaustão: efeitos e fundamentos da regulação de atividades com dados pessoais

Juiz de Fora/MG
2022

Eduardo Khoury Alves

Privacidade à exaustão: efeitos e fundamentos da regulação de atividades com dados pessoais

Dissertação apresentada ao Programa de Pós-Graduação *Stricto Sensu* da Faculdade de Direito da Universidade Federal de Juiz de Fora como requisito parcial para a obtenção do título de Mestre em Direito na área de concentração Direito e Inovação, sob a orientação do Prof. Dr. Sérgio Marcos Carvalho de Ávila Negri.

Juiz de Fora/MG
2022

Ficha catalográfica elaborada através do programa de geração
automática da Biblioteca Universitária da UFJF,
com os dados fornecidos pelo(a) autor(a)

Khoury Alves, Eduardo.
Privacidade à exaustão : efeitos e fundamentos da regulação de atividades com dados pessoais / Eduardo Khoury Alves. -- 2022.
171 f.

Orientador: Sérgio Marcos Carvalho de Ávila Negri
Dissertação (mestrado acadêmico) - Universidade Federal de Juiz de Fora, Faculdade de Direito. Programa de Pós-Graduação em Direito, 2022.

1. proteção de dados pessoais. 2. privacidade. 3. regulação. 4. concorrência. 5. direito comparado. I. Negri, Sérgio Marcos Carvalho de Ávila, orient. II. Título.

Eduardo Khoury Alves

Privacidade à exaustão: efeitos e fundamentos da regulação de atividades com dados pessoais

Dissertação apresentada ao Programa de Pós-graduação em Direito da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Mestre em Direito. Área de concentração: Direito e Inovação

Aprovada em 09 de dezembro de 2022.

BANCA EXAMINADORA

Prof. Dr. Sergio Marcos Carvalho de Ávila Negri - Orientador e Presidente da Banca
Universidade Federal de Juiz de Fora

Profa. Dra. Claudia Maria Toledo da Silveira - Membro titular interno
Universidade Federal de Juiz de Fora

Prof. Dr. Renato César Cardoso - Membro titular externo
Universidade Federal de Minas Gerais

Juiz de Fora, 28/11/2022.

Documento assinado eletronicamente por **Sergio Marcos Carvalho de Avila Negri**,

Professor(a), em 13/12/2022, às 23:33, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Claudia Maria Toledo da Silveira**, **Professor(a)**, em 15/12/2022, às 00:12, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Renato César Cardoso**, **Usuário Externo**, em 15/12/2022, às 09:59, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no Portal do SEI-Ufjf (www2.ufjf.br/SEI) através do ícone Conferência de Documentos, informando o código verificador **1056223** e o código CRC **7180CCC6**.

Para Virgínia.

RESUMO

O marco normativo brasileiro sobre proteção de dados pessoais decorreu de importantes debates, e predominou o referencial normativo-teórico que embasa o Regulamento Geral de Proteção de Dados Europeu. Esta influência é marcante na definição do conceito de dado pessoal e na sua abertura semântica, e no erguimento de uma normativa que se calca em uma miríade de objetivos fundamentais. Ao declarar como objetivos e fundamentos tantos tópicos correspondentes a direitos de natureza fundamental-constitucional, a LGPD atrai para si a necessidade de harmonizá-los, preservando e aprimorando os espaços de liberdade que visam a tutelar. Tendo isto em vista, colocou-se como problema avaliar se, e em que medida, o regime jurídico estatuído pela LGPD é capaz de promover uma harmonização desejável entre estes objetivos e fundamentos. Para tanto, a hipótese deste trabalho, relacionada ao peso desproporcional do direito à privacidade na construção teórico-normativa que fundamenta a LGPD, foi construída a partir da teoria enraizada, e partiu de um estudo fenomenológico das interações que produzem dados pessoais e da forma de utilização destes dados. O objetivo é responder em que se fundamenta a intervenção jurídica sobre atividades com dados pessoais e quais são os efeitos potenciais desta intervenção, de forma a melhor compreender como o direito pode atuar para atingir os objetivos constitucionais em um cenário socioeconômico em que atividades que dependem de dados, pessoais ou não, adquirem inédita importância. Inicialmente, propôs-se uma breve discussão acerca do papel do direito no estabelecimento e modificação das relações sociais, a partir de um olhar da interação com outras disciplinas e atividades, seguida de uma análise da forma com que se estruturam as principais teorias que embasam estudos sociológicos e proposições jurídicas relacionadas a dados pessoais. Em seguida, analisou-se a estruturação do conceito de dado pessoal e do conceito de informação que o integra, e a sua inserção enquanto categoria jurídica e as suas funções no sistema normativo sobre dados pessoais. Avança-se para análise a respeito da relação individual com as possibilidades de controle da informação, e da maneira com que determinados elementos favorecem ou desfavorecem a existência e a concretização destas possibilidades. Por fim, o estudo dedica-se à análise da interação da regulação sobre atividades com dados pessoais e os mercados de atividades que dependem de dados pessoais, com destaque para estudos empíricos realizados no continente europeu que têm sugerido maior concentração de mercado e outras disfunções em decorrência da atual disciplina de proteção de dados pessoais. É possível concluir que os referenciais teóricos da disciplina jurídica de proteção de dados pessoais adotada no

Brasil conduzem a uma intervenção jurídica pouco capaz de dialogar com práticas contemporâneas que se utilizam de dados, com resultados pouco satisfatórios para titulares de dados e para agentes de tratamento, assim como para a disciplina, que carece de coerência condizente com a sua abrangência e impacto. As conclusões do estudo enumeram frentes de pesquisa importantes para o aprimoramento da disciplina de proteção de dados pessoais, dos quais se destacam a investigação do potencial da discursividade multimodal para propiciar melhores condições de escolha a titulares de dados pessoais; a possibilidade de estruturação de arranjos normativos em que titulares de dados detenham maior poder sobre a coleta e utilização de dados pessoais; a viabilização de estruturas jurídico-econômicas que favoreçam a desconcentração de poder econômico no mercado de tecnologia.

Palavras-chave: Proteção de Dados Pessoais. Privacidade. Regulação. Concorrência. Direito Comparado.

ABSTRACT

The Brazilian data protection law is strongly influenced by the theoretical and normative perspectives that underlie the European General Data Protection Regulation. This influence can be perceived in the definition of personal data and in the openness of the semantic content of this concept, as well as in the fact that both norms are structured upon a set of principles. The Brazilian General Data Protection Regulation (LGPD) enunciates to be founded upon and aimed to protect several fundamental principles, which means it must fulfill the task of balancing them by protecting the different dimensions of freedom they encompass. This study aims at evaluating if, and to which extent, LGPD is capable of promoting balance between those principles in a desirable form. In order to accomplish this, grounded theory has been implemented, in conjunction with a phenomenological approach, to formulate the hypothesis that the right to privacy is disproportionately embedded in the core foundations of LGPD. This has led to an analysis which involves taking into consideration the foundations and objectives of data protection regulation, as well as its effects, starting from the observation of the phenomena it is intended to work upon. The objective of this approach is to better apprehend how the data protection regulation framework might work in order to achieve constitutional objectives and promote an adequate balance amongst them, considering the major importance of data treatment activities in contemporary economy. The study starts with a short discussion on the role of Law in establishing and altering social relations, especially in light of the possible interactions between the Law and other fields of knowledge. Some of the most common ways of conceptualizing data protection problems and regulation are also part of the discussion. The study then analyses the concepts of personal data and of information, as well as its functions in the regulatory framework, and moves to evaluate subjective perspectives of control information and how data protection regulation might or might not work to provide control in different stages of information use. At last, some of the implications of data protection regulation to personal data markets are exposed and discussed, especially in light of discoveries that show increased market concentration as a result of such regulation. It was possible to conclude that the theoretical foundations of Brazilian data protection regulation lead to a kind of intervention that is incapable of effectively approaching and dialoguing with contemporary data treatment activities, which results in poor outcomes for data subjects and treatment agents, as well as for the consistency of data protection theory regulation. The conclusions leave several perspectives open for discussion and research, such as investigations regarding the potentialities of

multimodal argumentation in providing better choice-conditions for data subjects; possibilities regarding the structuring of arrangements that provide more meaningful power to data subjects, especially in the later stages of information use; the considering of regulatory and economic structures that favor power dispersion in technology and data markets.

Keywords: Data Protection. Privacy. Regulation. Concurrency. Compared Law.

SUMÁRIO

INTRODUÇÃO

A Lei 13.709 de 14 de agosto de 2018 (Lei Geral de Proteção de Dados Pessoais – LGPD) traz, em seus artigos 1º e 2º, uma série de objetivos e fundamentos, entre os quais a autodeterminação informativa, o respeito à privacidade, a liberdade de expressão, o desenvolvimento tecnológico, a livre iniciativa e a livre concorrência. Ao declarar como objetivos e fundamentos tantos tópicos correspondentes a direitos de natureza fundamental-constitucional, a LGPD atrai para si a necessidade de harmonizá-los, preservando e aprimorando os espaços de liberdade tutelados por cada um. Tendo isto em vista, colocou-se como problema avaliar se, e em que medida, o regime jurídico estatuído pela LGPD é capaz de promover uma harmonização desejável entre estes objetivos e fundamentos. Para tanto, o estudo foi voltado, inicialmente, a apreender o objeto da regulação, e, em seguida, os fundamentos para a intervenção jurídica sobre ele.

Com isso, este trabalho trata essencialmente dos fundamentos e efeitos da intervenção jurídica que recai sobre atividades com dados pessoais, e é calcado em abordagem exploratória e fenomenológica. A abordagem deste estudo é em torno das *próprias atividades* que requerem intervenção jurídica, para, a partir de suas características intrínsecas e contextuais, apreender o tipo de problema que colocam, e, com isso, os fundamentos e os efeitos da intervenção jurídica efetiva ou desejada sobre estas atividades. Esta abordagem colocou no centro as atividades com dados pessoais que ocorrem por meio de tecnologias da informação e comunicação, que, ao mesmo tempo, representam a maior parte destas atividades e são aquelas mais afetadas pela legislação em matéria de proteção de dados pessoais. Quanto a este contexto, apesar de o cenário atual, dominado pelos tecnopólios e agências de vigilância governamental, inspirar certo pessimismo quanto aos mercados de dados, as tecnologias da informação apresentam enorme potencial para o aprimoramento democrático, bem como para a inovação pulverizada e acessível técnica e economicamente e para avanços importantes nas mais diversas disciplinas científicas, pelo que o adequado equilíbrio na regulamentação de atividades com dados pessoais é fundamental para harmonizar objetivos constitucionais e preservar espaços de liberdade. Todos estes elementos justificam as opções metodológicas adotadas e o objeto de pesquisa selecionado.

Esta abordagem se mostrou necessária durante o processo iterativo de pesquisa que lhe deu origem e por meio do qual se desenvolveu este estudo, com fundamento na fenomenologia e na teoria referenciada. Especificamente, a partir de metodologia exploratória e

fenomenológica, a revisão bibliográfica de estudos empíricos voltados a avaliar eventuais impactos do GDPR no ambiente concorrencial das jurisdições em que é aplicável tornou claro que alguns dos efeitos adversos da disciplina recaíam sobre aspectos que deveria tutelar ou mesmo que eram parte de seu arcabouço fundamental. Por outro lado, observou-se que estes efeitos decorrem de um modelo jurídico que foi replicado em outras jurisdições, como o Brasil, e que representou uma estabilização epistemológica das discussões acerca de proteção de dados pessoais em torno do direito à privacidade. Este aspecto foi colocado lado a lado ao problema inicialmente proposto quanto à harmonização entre os fundamentos da LGPD, e originou a hipótese de que uma inadequada construção conceitual-normativa decorrente do conceito de privacidade teria o condão de causar desequilíbrios em função da regulação.

Especificamente, sem prejuízo de análises mais detidas a serem levadas a cabo adiante, pode-se dizer que a tensão entre "privacidade" e "autodeterminação informativa", por um lado, e princípios como a promoção da inovação, da concorrência e do progresso científico, de outro, não é usualmente capturada adequadamente com relação aos seus pressupostos e a seus efeitos, levando a uma predisposição à proteção de valores jurídicos abstratos, sem que o exercício de ponderação recaia sobre a eficácia de tais medidas protetivas, i.e., ao fato de atingir as metas previstas, ou sobre a sua adequação, i.e., a sua aptidão para, em abstrato, atingir tais objetivos. Em outras palavras, os resultados da normativa têm demonstrado efeitos adversos em alguns campos, como no ambiente concorrencial, e ineficientes nos aspectos sobre os quais teria maior poder de tutela, como na proteção à autodeterminação informativa. Adicionalmente, leva a um subaproveitamento do potencial de inovações tecnológicas que dependem de dados.

Para um país como o Brasil, que caminha para amadurecer a sua disciplina jurídica de proteção de dados pessoais, este cenário representa desafios e oportunidades, enunciando caminhos para uma legislação (e sua regulação e aplicação) que atinja, na maior medida possível, os objetivos e fundamentos constitucionais que enuncia. A confiança é elemento primordial nas relações que empreendemos e que geram dados que alimentam estes sistemas. Ao estarmos em constante contato com sensores conectados, passamos a fazer parte de vários ambientes que não o construído que enxergamos; essa afirmação não é metafísica, mas, ao contrário, procura nos fazer visualizar o fato de que os ambientes construídos, naturais e os objetos que utilizamos cotidianamente estão permeados por sensores que apreendem informações e as transformam em dados. Neste constante contato, fatos do mundo, inclusive da interação das pessoas com outras pessoas ou com o ambiente, geram informações que são coletadas, armazenadas e processadas de diversas maneiras, inclusive gerando mais

informações a partir de sua combinação, de inferências, e de técnicas inúmeras de análise de dados, que ultimamente permitem a geração de conhecimento, mais ou menos valioso, e com potencialidades diversas para aqueles que sobre eles detenham controle e tenham a capacidade de utilizá-lo. Posteriormente, esses conjuntos de dados têm o condão de impactar em diversas searas das vidas daqueles a que se lhe referem.

Não nos apercebermos, ou nos apercebermos parcialmente enquanto parte dessa arquitetura é crucial para que se mantenha o ciclo de vulnerabilidade e torna inimaginável a privacidade como direito de manter o controle sobre as próprias informações, caracterizando um *déficit perceptivo* que é representativo da opacidade e invisibilidade dos ambientes virtuais responsáveis por atividades de tratamento de dados. Isto posto, a fenomenologia é relevante para a estruturação deste trabalho sob dois aspectos. Inicialmente, enquanto metodologia que permitiu, a partir do problema inicialmente proposto, a delimitação do objeto de pesquisa e o seu desdobramento, que, em associação ao método exploratório e se valendo da teoria enraizada (GLASER et al., 1967), levou à formulação e desenvolvimento das hipóteses ora apresentadas. Por outro lado, e neste ponto especialmente a partir da fenomenologia da percepção (MERLEAU-PONTY, 1954), serve como lente para enxergar aspectos fundamentais do que ora se coloca sob análise, a saber, os próprios fenômenos vislumbrados no contexto da análise de atividades de tratamento de dados pessoais.

Em apertada síntese, cabe mencionar que a fenomenologia visa à superação de um dito "dualismo cartesiano" entre mundo objetivo e mundo subjetivo; entre intelectualismo e empirismo. Trata a experiência como algo de que o indivíduo é parte, e a apreensão da realidade como a aproximação, pelo sujeito, das constantes perceptivas da coisa através de um ponto de vista, que, outrossim, não corresponde a uma aparência privilegiada, mas à "armação de relações às quais todas as aparências satisfazem". Compreende que o pensamento objetivo predominante reduz os fenômenos que ligam o sujeito ao mundo ao substituí-los por conceituações do objeto em si e do sujeito como pura consciência, que acaba por fazer subsistir apenas as qualidades imediatamente sensíveis dos objetos e do meio, dentre as quais predominam as qualidades visuais, que têm uma aparência de autonomia e "antes nos apresentam um objeto do que nos introduzem em uma atmosfera" (MERLEAU-PONTY, 1954).

O raciocínio estritamente objetivo, conforme a clássica divisão cartesiana, parece dificultar a compreensão das arquiteturas digitais, uma vez que elas se sobrepõem ao ambiente físico e o permeiam de forma predominantemente invisível, porém absolutamente real, por meio

de relações que nos inserem em diversos ambientes que não enxergamos. Por outro lado, na interação com objetos e com o meio dotados de sensores e conectados à internet, é preciso que as pessoas sejam, ao menos na maior parte dos casos, capazes de compreender as potencialidades reais destes objetos, que, sobretudo a partir da incorporação de funcionalidades de conectividade, de capacidade de apreender aspectos do ambiente, i,e., sensores, como microfones, câmeras, termômetros, giroscópios etc., e inteligência artificial, e de algoritmos capazes de interpretar esta informação, dota uma crescente parcela dos objetos com que lidamos e ambientes em que vivemos de capacidades que diferem daquelas dos objetos em si, não fossem estes dotados de determinados atributos relacionados às tecnologias da informação.

Como um exemplo que se pretende bastante simples – e com as escusas do leitor familiarizado com o assunto pelo raciocínio simplório –, falemos sobre um termômetro digital, daqueles que se coloca sob a axila. Um termômetro comum, ainda que digital, porém sem qualquer função de conectividade, apenas tem a capacidade relevante de viabilizar ao usuário aferir a própria temperatura. Um termômetro, por outro lado, que incorpore conectividade por WiFi, ou Bluetooth, e envie para o *smartphone* do usuário a temperatura aferida, também ostenta esta capacidade; para além disso, provavelmente ostenta diversas outras, como, por exemplo, a de enviar para o controlador do serviço informações como a temperatura aferida pelo usuário, e, possivelmente, outras, como o horário em que foi aferida a temperatura e o local em que se supõe estava o usuário no momento da aferição (a partir da localização de seu *smartphone*). Para além disso, podemos imaginar um termômetro que seja capaz de informar ao usuário se, ao colocá-lo sob a axila, este está posicionado corretamente (i.e., caso seja preciso que seja posicionado na horizontal, e o usuário o coloque em ângulo de 45° com relação à axila, será enviada notificação para o *smartphone* do usuário), o que se afigura útil para assegurar maior precisão na aferição da temperatura. Para desempenhar esta funcionalidade, o termômetro deverá ter algum tipo de sensor, como um giroscópio, que é capaz de "dizer ao aparelho", ou seja, transmitir ao *chip* do termômetro a informação que pode ser interpretada como a posição do termômetro. A partir disso, o termômetro também poderá informar ao controlador do serviço informações sobre a posição do termômetro, e possivelmente acerca da forma como o usuário utiliza o objeto (i.e., se o movimenta rapidamente ou devagar; se segura o termômetro com força etc.), o que pode permitir inferências sobre o estado do usuário (i.e., nervoso, calmo etc.).

A partir deste exemplo, queremos chamar atenção para o fato de que, contemporaneamente, os objetos e o ambiente estão permeados destes sensores e de

mecanismos para o armazenamento, processamento e transferência das informações por eles coletadas, que, por sua vez, frequentemente dão origem a mais informações (como, por exemplo, no caso acima, informações objetivas a respeito do uso do termômetro podem viabilizar inferências, mais ou menos apuradas, sobre aspectos como o estado mental do usuário). Assim, estes objetos e sensores ostentam muito mais potencialidades do que aparentam – não em razão de capacidades sobrenaturais ou místicas, mas porque associamos a sua forma a determinadas funcionalidades, ao passo que, devido ao embutimento de *chips* e sensores, estas funcionalidades são ampliadas de maneira que muitas vezes não pode ser apreendida a partir de um olhar do objeto externamente, ao menos sem que se tenha um conhecimento prévio acerca destas potencialidades.

Mearleau-Ponty (1954), ao descrever a atividade de percepção visual como forma de explicitar a sua tese, retoma a conhecida experiência de se colocar um anteparo entre o observador e a coisa observada, notando que "o fator decisivo no fenômeno da constância, que o anteparo põe fora de jogo e que funciona na visão livre, é a articulação do conjunto do campo, a riqueza e a sutileza das estruturas que ele comporta" (MERLEAU-PONTY, 1954, p. 413). Desta forma, ao olhar um cenário através de um anteparo, o observador não é capaz de "dominar" (*übershcauen*) as relações de iluminação e perceber no espaço visível as subordinações entre os objetos a partir da apreensão de suas características próprias de claridade. Poderemos caracterizar os ambientes digitais, e.g. sistemas operacionais, smartphones, objetos inteligentes em geral, lojas e outros ambientes com sensores biométricos etc. como *anteparos*, que frequentemente impedem que enxerguemos os próprios sensores e/ou os contextos em que estão inseridos, i.e., redes com que se conectam e compartilham dados, e ainda as maneiras com que são utilizados estes dados em inúmeras práticas sociais e econômicas relevantes. A nossa interação com os ambientes digitais se baseia em uma série de relações de confiança, decorrentes da necessária aceitação das diversas camadas de *hardware* e *software* que operam para o seu funcionamento, e que conduzem as nossas experiências – experiências que se visa regular mediante a disciplina jurídica de proteção de dados pessoais, assim como os seus produtos.

Lawrence Lessig (1996) escreveu, sob a ótica do direito norte-americano, acerca do asseguramento do âmbito de proteção de direitos constitucionais diante de mudanças tecnológicas. Após analisar algumas decisões da Suprema Corte daquele país, conclui que "Quando as tecnologias do mundo [em que foi criada a Constituição] mudam, somos confrontados com uma escolha. (...). Estas são nossas escolhas democráticas, e são escolhas

reais" (LESSIG, 1996, p. 43). O problema proposto, acerca da eventual harmonização dos espaços de liberdade fundamental elencados na LGPD como objetivos e fundamentos, representa, diante da ubiquidade de atividades com dados na economia contemporânea, uma importante fronteira entre o desmoronamento de espaços de liberdade ou sua renovação diante de uma realidade modificada pelas tecnologias da informação e comunicação e as novas potencialidades embutidas em objetos e espaços – mas sempre operacionalizadas por outro agentes.

No capítulo 1, trata-se da conceituação de *complexidade organizada* e das razões pelas quais é preciso compreender os potenciais efeitos da intervenção sobre elementos de sistemas complexamente organizados, notavelmente o potencial de causar problemas não diretamente associáveis à intervenção. O capítulo 2 deste trabalho abriga a maior parte das discussões a respeito das incongruências entre o regime jurídico de proteção de dados pessoais e a materialidade sobre a qual objetiva atuar, enquanto os capítulos 3 e 4 tratam dos efeitos desta incongruência sobre as possibilidades de controle individual sobre a informação e sobre as atividades econômicas que dependem de dados pessoais, respectivamente.

As limitações deste trabalho se relacionam sobretudo à sua abertura na delimitação do problema investigado e na amplitude requerida pela abordagem fenomenológica. Assim, não foram empreendidas análises ou coletados dados empíricos sobre aspectos específicos da regulação em matéria de dados pessoais no Brasil, tampouco avaliados os recentes esforços da Autoridade Nacional de Proteção de Dados Pessoais em orientar agentes de tratamento de dados e regulamentar aspectos da LGPD. Estas limitações, por outro lado, denotam a abertura de significativos espaços de pesquisa e desenvolvimento, capazes de responder a questionamentos e a necessidades específicos a partir de um arcabouço teórico consistente e que comporta o desenvolvimento iterativo da disciplina jurídica de proteção de dados pessoais.

1 COMPLEXIDADE, INTERVENÇÃO E NARRATIVAS

Neste primeiro capítulo, propomos, inicialmente, uma abordagem que esperamos ser capaz de se colocar presente durante a leitura deste trabalho, com o objetivo de propiciar um olhar sobre a regulação jurídica que privilegie, por um lado, o atingimento de resultados desejáveis por meio da regulação, e, por outro, a coerência do próprio ordenamento jurídico.

1.1 Complexidade e intervenção

Esta seção tratará do conceito de complexidade organizada e suas particularidades, especialmente naquilo que se relaciona com alguma espécie de intervenção em sistemas complexamente organizados.

1.1.1 Sistemas complexos e reduções equivocadas

A utilização de analogias entre espaços virtuais e físicos, ou entre a *arquitetura* da *internet* e a arquitetura dos ambientes em que vivemos, é relativamente comum ao tratar da questão da regulação da rede. Uma das mais relevantes contribuições para o tema que se utilizou desta analogia foi a de Lessig (1999), no que foi seguido por Solove (2006), em que o intrínseco caráter normativo do *hardware* e *software* que compõem a *internet* desempenha papel central em quaisquer experiências que ocorram *por meio* desta rede. Lessig (1999) discute as formas com que esta *regulação* exercida pela arquitetura sobre o comportamento daqueles que dela se utilizam interage com outros elementos com caráter regulador, inclusive as normas jurídicas. Acresce-se a isto o fato de que os efeitos de determinado regramento jurídico, por intervirem em sistemas complexos, podem exercer efeitos imprevistos e mesmo antagônicos com relação ao objetivo perseguido. Naturalmente, empresas como o Google e Facebook também descobriram há algum tempo o potencial normativo dos ambientes que estruturam, o que, associado ao enorme conjunto de conhecimento que são capazes de sintetizar gera inúmeras possibilidades de atividades lucrativas e de sua perpetuação. Percebamos que o potencial lesivo não está propriamente na arquitetura erigida, mas na forma como ela promove possibilidades para a criação e exploração de vulnerabilidades, conforme distinção proposta por Calo (2017). Aparentemente, este é o problema que devemos nos esforçar para compreender e endereçar: o porquê de estas possibilidades de vulnerabilização e exploração serem tão abundantes, e como superá-las.

O valor que se pode extrair das analogias entre a regulação da rede e a regulação do espaço físico, por outro lado, pode ir além. Em 1961, a jornalista estado-unidense Jane Jacobs (1961) publicou uma das obras mais influentes do urbanismo, cujo nome em português é “Morte

e Vida de Grandes Cidades". Jacobs (1961, p. 13) discute, sobretudo, "o *tipo* de problema que as cidades apresentam – um problema de manejar a complexidade ordenada". Afinal, organizar uma cidade não se trata de propiciar um trânsito fluido de automóveis ou impedir que crimes ocorram – se trata de propiciar que inúmeras atividades relevantes ocorram, ao passo que se evita a erosão da própria cidade, que ultimamente é o espaço onde toda a vida social ocorre. Ou seja: trata-se de propiciar um trânsito fluido de automóveis, e de impedir que crimes ocorram; mas os objetivos não são atingidos quando isso ocorre *ao custo* de outras atividades que deveriam encontrar terreno fértil nas cidades para o seu desenvolvimento, ou quando as medidas adotadas destroem os próprios elementos que viabilizam a existência da cidade e o desenvolvimento de atividades humanas em seu seio.

Cass Sunstein (2018, p. 20), ao identificar a obra de Jacobs (1961) como "uma inspiração central" para seu livro acerca da democracia no contexto das mídias sociais, expôs aquilo que Jacobs identificou como elementos propulsores de *vida* das cidades a partir do termo *serendipity*, descrevendo a qualidade de uma estrutura que propicie *escolhas* pelos indivíduos que dela se utilizam para desenvolver suas atividades; em que "as pessoas frequentemente encontrem pontos de vista e tópicos que não selecionaram especificamente. Este tipo de estrutura é, na verdade, uma arquitetura de escolha de que indivíduos e grupos se beneficiam enormemente" (SUNSTEIN, 2018, p. 11). Os espaços públicos compreendem aqueles em que os cidadãos *compartilham* experiências, e que representa uma espécie de "cola social" (SUNSTEIN, 2018, p 11). No contexto do debate sobre a democracia nos ambientes de rede, Sunstein (2018) reflete sobre estas duas precondições para uma democracia funcionante: (1) encontros fortuitos; e (2) experiências compartilhadas, ambos dependentes das condições de existência de diversidade.

Um dos fundamentos do planejamento urbano modernista consiste na concepção de que a cidade deve ser organizada de acordo com uma setorização de seus usos, i.e., zona residencial, zona de comércio etc. Ao longo dos anos, a noção mostrou-se absolutamente equivocada, o que foi habilmente demonstrado por Jacobs (1961). Um exemplo marcante é o caso do distrito de Boston chamado North End, que é uma área tradicional, de baixa renda, então misturada à indústria pesada, com a maior densidade habitacional da cidade e considerada uma "zona de cortiços". Jacobs (1961) nota que, ao visitar o North End em 1959, teve uma impressão muito diversa daquela que tivera vinte anos antes: muitos prédios haviam sido reformados, casas haviam sido ampliadas, havia diversidade de comércio e muitas pessoas nas ruas. A taxa de mortalidade infantil era menor que a média municipal e o índice de mortes por tuberculose

baixo para a época. No entanto, os responsáveis por formular políticas públicas e os bancos continuavam a considerar o local uma "zona de cortiços". A investigação de Jacobs (1961) trata tanto de compreender os equívocos que transformaram muitos bairros e cidades em verdadeiros fracassos (inclusive muitos que eram vistos como exemplares), quanto em destacar que elementos propiciaram avanços como o que vislumbrou no North End, a despeito da descrença e falta de apoio por parte dos formuladores de políticas públicas e investidores. Especialmente, busca compreender por que as teorias do urbanismo moderno não eram capazes de enxergar as virtudes de locais como o *North End*. Urbanistas, legisladores e banqueiros apenas estavam aplicando a teoria que haviam aprendido (Jacobs, 1961, p. 10). O que suas lentes os propiciavam enxergar? O que elas *não* propiciavam enxergar? Como isso impactou na formulação de políticas públicas e nos seus resultados? Como isso se refere aos objetivos perseguidos pelas políticas urbanas? Identificar que as cidades representavam problemas do tipo *complexidade organizada* foi central para que Jacobs pudesse responder a estas e outras questões importantes.

A *complexidade organizada* se refere a "problemas que envolvem uma abordagem simultânea de um número mensurável de fatores inter-relacionados num todo orgânico" (JACOBS, 1961, p. 481). A compreensão do papel da *inter-relação* das variáveis representa a questão central, em contraposição ao mero número de variáveis e sua distribuição estatística, que corresponderia a problemas de complexidade *desorganizada*. As ciências biológicas, desde a década de 1930, quando se observou o limiar do desenvolvimento de métodos analíticos efetivos para abordar a complexidade organizada, vislumbraram enorme progresso, que representa um ganho social sentido de forma generalizada. Basta observarmos que no início do século XX a sangria era prescrita pelo *establishment* médico – e esta representa apenas uma das práticas *deletérias* para a saúde realizadas em nome da medicina, ainda hodiernamente, porém em proporções absolutas antes da adoção de métodos científicos *consistentes com o tipo de problema* que este campo do conhecimento se propõe a resolver. Este progresso das ciências biológicas revela algo importante sobre outros problemas de complexidade organizada: que problemas desse *tipo* podem ser analisados – "que só cabe encará-los como passíveis de compreensão, em vez de considerá-los 'sinistra e fatidicamente irracionais'" (JACOBS, 1961, p. 481).

Afirmar que dois campos apresentam problemas de *tipos similares* é bastante diverso de afirmar que os problemas seriam os mesmos. As táticas para compreendê-los, outrossim, podem ser similares. Jacobs (1961) destacou os modos de reflexão que reputou mais importantes ao lidarmos com *este tipo* de problema: (1) refletir sobre os processos; (2) usar de

indução, raciocinando do particular para o genérico, em vez do contrário; (3) procurar indícios 'não-médios' que envolvam uma quantidade bem pequena de coisas, as quais revelem como funciona uma quantidade maior e 'média'. O que apreendemos disto é a relevância de abordar o problema a partir dos elementos sobre os quais é instado a atuar, o que envolve o próprio objeto da regulação e os problemas que se deve endereçar.

1.1.2 Interação regulatória e incentivos

Em *Code*, o constitucionalista Lawrence Lessig (1999) inicia a discussão sobre o *objeto da regulação* a partir de menção aos escritos do inglês John Stuart Mill. Mill (1859), na obra *On Liberty*, desenvolveu sua noção de liberdade a partir da compreensão de que devem ser suprimidas as forças que a ameaçam – que podem, ou não, terem como fonte os governos. O método de Mill questiona, primeiramente, com qual ameaça à liberdade lidamos; e, então, como podemos a ela resistir. O poder deste método é nos remeter à necessidade de focar nas ameaças à liberdade que existem em momento e local específicos, e não meramente em abstrato (LESSIG, 1999). Deixar de compreender quais *reguladores* estão presentes no fenômeno que se deseja regular, e de como interagem com o Direito, reduz drasticamente as chances de que se regule eficazmente. Isto representa um risco para a ocorrência de *excessividade* de regulação, de maneira contrária aos valores fundamentais de liberdade. Este risco tem se mostrado cada vez mais concreto devido à concentração de mercado marcante no setor de tecnologia e ao caráter impositivo da sua arquitetura.

Lessig (1999) nos convida a pensar acerca dessas diversas "forças", ou limitações, que influenciam as escolhas e o comportamento humanos. Em dados momento e espaço, atuam sobre aquele que tem a sua conduta *regulada* diversos limitadores, aos quais a soma, em uma operação que pode ser analogicamente considerada vetorial, corresponde à "regulação" existente sobre esta conduta. Assim, neste modelo, "regulação" do comportamento não equivale a regras jurídicas *apenas*, mas a uma espécie de sistema de incentivos. Muitos viram os argumentos de Lessig (1999) como argumentos pela redução do papel do Estado na regulação das interações pela rede, o que não poderia estar mais distante da realidade. O constitucionalista não busca, com sua obra, dissertar sobre o que é o Direito ou quais são as suas fontes, tampouco busca equiparar o que chama de *modalidades de regulação* às normas jurídicas. O que ele chama de *modalidades de regulação* nada mais são do que elementos que influenciam a conduta

humana na realidade. Sobre o sujeito instado a agir atuam diversos limitadores, ou incentivos, para se conduzir desta ou daquela maneira. O somatório destas forças exercerá determinado grau de influência sobre o sujeito, a que Lessig (1999) atribuiu caráter regulatório. O Direito é um destes elementos, e interage com os demais, pelo mero fato de coexistirem. O Direito compete com diversos incentivos à conduta humana, e é frequentemente sobrepujado quando não se presta a estabilizar os conflitos sociais que almeja regular.

Contemporaneamente a estes debates, em 1997, Joel Reidenberg (1997) publicou o seminal artigo *Lex informatica: The formulation of Information Policy Rules through Technology*, em que identificou o tratamento do conteúdo, o tratamento das informações pessoais, e a preservação dos direitos de autor, como três áreas então em estado crítico diante no ambiente da rede, e que o acesso, a distribuição e o uso da informação moldariam a confiança e a justiça no século XXI para cidadãos, negócios e governos (REIDENBERG, 1997, pp. 554-555). Ele argumentou que em ambientes de rede e na sociedade da informação o Direito e o Estado não seriam a única fonte normativa, e que as "capacidades tecnológicas e escolhas de design impõem regras aos participantes", o que chamou de "*Lex Informatica*", argumentando que os responsáveis por políticas públicas deveriam "conhecê-la, conscientemente reconhecê-la, e encorajá-la", especialmente de forma a que "capacidades e funções técnicas específicas incorporem objetivos públicos" (REIDENBERG, 1997, p. 554).

Ao destacar que o direito representa um dentre diversos *reguladores* da conduta humana que ocorrem na realidade, Lessig (1999) demonstrou, por um lado, que algumas formas de regulação são *mais eficazes* do que a regulação pelo direito. Por outro, que *maior eficácia* regulatória não representa por si maior atingimento dos valores da sociedade – ou, simplificadamente, os valores representados por um regime jurídico específico, ou aqueles constitucionalmente protegidos. Ao deixar de regular, ou ao regular ineficaz ou inadequadamente, o Direito cede espaço para reguladores como os mercados, as normas sociais e as arquiteturas, físicas ou digitais. Conforme observado por Cardoso et al. (2018), o momento é de esgotamento do modelo de segregação estrita entre as ciências, que, isoladamente, produziram enorme conhecimento especializado, que ora precisa ser integrado e harmonizado para o atingimento de resultados relevantes.

É frequente pensar nos mercados, normas sociais e arquitetura como dados, ou seja, imutáveis, enquanto apenas o Direito seria moldável. Porém, as relações entre estes reguladores são complexas, e se afetam mutuamente de diversas maneiras. Regulação eficiente com relação

aos fins constitucionalmente perseguidos deve levar em conta os efeitos que deseja atingir, e compreender como o direito e os demais reguladores interagem. Conforme observa Reidenberg (1997), "Culturalmente, engenheiros começam a projetar quando são defrontados com objetivos específicos" (REIDENBERG, 1997, p. 592). Também a este respeito, mais recentemente, Sunstein (2017) observou que as pessoas responsáveis pela organização de *arquiteturas de escolha* sempre terão o papel de conduzir as experiências que se realizam no meio que projetam, ainda que não tenham esta intencionalidade. A interação entre as diversas modalidades de regulação é dinâmica, e o projeto de políticas públicas requer a consideração das modificações normativas que se deseja realizar, e também a previsão dos efeitos *responsivos* que estas mudanças estimularão (LESSIG, 1999). Esta previsibilidade, por outro lado, apresenta-se como outro problema central, embora essencial para a formulação e avaliação de políticas (CASA CIVIL DA PRESIDÊNCIA DA REPÚBLICA, 2018).

1.1.3 Decisão com base em evidências e Estado de Direito

Se a inserção de conhecimento científico no âmbito de políticas públicas tem o condão de trazer legitimidade para este processo, maior cuidado deve ser dedicado ao escrutínio das formas de conhecimento empregadas e da sua valoração no contexto em que ocorre. Conforme observou Jacobs (1961), o planejamento urbano moderno nos Estados Unidos foi marcado inicialmente pelo emprego de técnicas científicas voltadas para a análise de modelos de *complexidade desorganizada*, o que representou pouca diferença na concepção de cidade que se estava adotar e pouca capacidade explicativa dos problemas que visava a solucionar, mas teve a aptidão de legitimar relevantes decisões adotadas pelo seu caráter *científico*. No caso do direito, em que boa parte da literatura relacionada à tomada de decisão se volta a assegurar a sua legitimidade mediante o emprego da racionalidade (CARDOSO et al., 2018), a preocupação com este aspecto é ainda mais importante.

Beecher-Monas (2007) argumenta que o adequado funcionamento do sistema jurídico, a credibilidade do sistema judicial e a adequação das discussões jurídico-acadêmicas dependem da incorporação consistente do conhecimento científico atualizado produzido por diversas disciplinas. Por outro lado, o emprego de evidências científicas no contexto judicial não pode representar que se deixe a cargo da comunidade científica a decisão, tampouco que a valoração desta evidência deva transformar o juiz ou advogado em "cientista amador" (BEECHER-

MONAS, 2007, p. 7). Na verdade, o objeto da desmistificação do argumento científico e busca por sua acessibilidade aos atores responsáveis pela tomada de decisões é auxiliar na resolução de problemas jurídicos e de políticas públicas, especialmente a avaliação da validade de hipótese científica específica proferida por experts quanto à sua potencial utilidade para auxiliar na resolução de determinada controvérsia (BEECHER-MONAS, 2007).

É importante levar em consideração que, se por um lado, o Direito tem progressivamente atribuído voz à ciência em seus procedimentos, mediante o emprego do trabalho do cientista, "invariavelmente, a decisão é da autoridade decisória, que deve digerir a informação científica e entregar um produto (uma decisão judicial ou administrativa, uma lei, ou uma política pública)" (CARDOSO et al., 2018, pp. 118-119). Por esta razão, é importante que os encarregados de atuar no campo jurídico sejam capazes de compreender a produção científica a respeito daquilo sobre o que são instados a atuar: estes agentes não podem empregar os instrumentos jurídicos sem considerar o conhecimento disponível a respeito do seu objeto; não se pode, por outro lado, delegar a tarefa de decidir, tampouco deixar de lado o rigor técnico e as funções próprias do Direito. Introduzir o direito em um debate mais amplo, de aproximação entre as diferentes instâncias das Ciências, deverá ter o condão de levar a um aprimoramento das soluções jurídicas oferecidas pelas instituições (CARDOSO et al., 2018).

O argumento contra deixar a cargo de cientistas a decisão regulatória nos remete, ainda, a preocupações potencialmente mais relevantes no contexto dos ambientes de rede: no âmbito de relações mediadas pela tecnologia, o *hardware* e o *software* que compõem a rede e os terminais envolvidos na conexão desempenham papel regulador com características que o tornam ubíquo e pouco passível de superação. Assim, a inadequação das decisões de políticas públicas e judiciais neste âmbito pode deixar a cargo dos responsáveis por estes sistemas a regulação de parcela considerável das relações intermediadas pela rede – que, por sua vez, representam parcela cada vez maior de aspectos relevantes da vida social, política e econômica. O direito, assim, perde a sua relevância na tutela de âmbitos de liberdades fundamentais ao falhar em se inserir *adequadamente* neste contexto, o que constitui uma das hipóteses centrais deste trabalho.

No caso de políticas públicas e no âmbito judicial, caso desejemos tomar decisões baseadas em evidências – seja de forma prognóstica, seja para a avaliação de sistemas normativos ou *standards* decisórios correntes –, é preciso adotar padrões científicos consistentes (BEECHER-MONAS, 2007). Diante da impossibilidade, como regra, de

experimentação controlada nesta seara, e da complexidade inerente aos sistemas afetados pelo regramento jurídico, além da complexidade inerente ao próprio sistema jurídico, a única alternativa para a adoção consistente de pressupostos baseados em relações de causalidade é a sua determinação a partir de estudos baseados em dados observacionais. O estabelecimento deste tipo de relação, por sua vez, apresenta-se como extremamente desafiador (MOOJI et al., 2016), seja com relação ao estabelecimento de relações entre variáveis, seja para a sua significativa inserção no sistema em que se deseja intervir.

Se por um lado o discurso jurídico pode e deve ser amplamente melhorado a partir de seu embasamento no conhecimento científico produzido por diversos campos do conhecimento, é necessário que a abordagem para com tais estudos seja precedida de avaliação cuidadosa e de sua inserção em um procedimento lógico-racional que parta de ceticismo para com relação ao seu valor preditivo e explicativo da materialidade a que se refere. Embora estudos individualmente considerados constituam os blocos fundacionais do conhecimento científico, a sua confiabilidade está diretamente relacionada à existência de evidências convergentes decorrentes de diferentes tipos de estudos, experimentos e observações (BEECHER-MONAS, 2007, p. 180).

De acordo com o Avaliação de Políticas Públicas – Guia prático de análise ex ante (CASA CIVIL DA PRESIDÊNCIA REPÚBLICA et al., 2018, p. 12):

> [P]artindo-se da identificação e caracterização de um problema que demandaria intervenção do Estado, é necessário que sejam estabelecidos objetivos claros para a ação governamental, bem como um desenho que efetivamente permita alcançá-los, considerando, por exemplo, os incentivos aos agentes envolvidos. [...] A qualidade das informações obtidas e das decisões tomadas na análise ex ante afeta sobremaneira o desenvolvimento do ciclo da política pública [...] e independentemente da qualidade da análise e do planejamento realizado, na execução da política pública, deve-se monitorar e avaliar a intervenção para garantir que sejam alcançados os impactos esperados e planejados antes da implementação.

Na disciplina de proteção de dados e privacidade considerada amplamente, uma das consequências do emprego descontextualizado de conclusões de estudos empíricos, sobretudo da área comportamental, é vislumbrada na predominância entre estudiosos e formuladores de políticas públicas da crença na existência do que foi chamado de *paradoxo da privacidade*. Os estudos afetos à área da psicologia comportamental que embasam esta tese explicativa levam à adoção de posições a ela conformes. No entanto, estas posições e as soluções deles decorrentes apenas serão consistentes na medida em que a tese explicativa em que se fiam, de fato, explique

a realidade. Do contrário, corre-se o risco de que tais decisões não atinjam as finalidades desejadas a partir da adequada compreensão do problema, ou ainda, de que exerçam efeitos deletérios que sequer podem ser associados à intervenção em decorrência da ausência de uma teoria ou hipótese hábeis, o que a seu turno retira do campo de discussão aspectos essenciais do problema, perpetuando-o ou o exacerbando.

A explicitação destes elementos não pretende cobrir com manto de *necessária adequação* as medidas eventualmente empregadas, mas, ao contrário, visa a permitir que esta *potencial adequação* possa ser continuamente revisitada, de forma objetiva, e a partir da definição de critérios e indicadores que permitam observar os efeitos da intervenção. No âmbito da regulação da tecnologia e das legislações sobre proteção de dados pessoais, algumas abordagens acadêmicas têm se valido de metodologias e lentes que buscam apreender a complexidade característica do problema. Um elemento que tem se mostrado relevante é o recurso à ecologia. O trabalho de Calo (2017), que se fia na Teoria Ecológica da Percepção de Gibson (1979), e o de Prazeres (2021), que analisa o direito sob a ótica sistêmica, são exemplos notáveis. O que a ecologia tem propiciado é o olhar sobre o objeto de regulação a partir de seu contexto e a preocupação com a sua capacidade de permanência e de endereçamento dos problemas sobre os quais tem a aptidão de atuar, abrindo caminhos para uma regulação iterativa, capaz de desenvolver-se de acordo com os seus resultados e as modificações sofridas por seus objetos.

1.1.4 Iatrogenia e ecologia

Somente a organização do conhecimento, e a explicitação dos objetivos e fundamentos do processo intelectual podem levar a decisões potencialmente efetivas e passíveis de escrutínio e aprimoramento. Decisões irrefletidas (ou decisões que não permitem escrutínio), por outro lado, costumeiramente são mais do que inócuas ou ilegítimas – elas atingem efeitos diversos daqueles explicitados e/ou intendidos. A *intervenção ingênua*, ou seja, a intervenção em um sistema que deixa de dar conta da interação entre as suas variáveis, frequentemente dá causa a uma série de danos ao sistema, que geralmente permanecem escondidos ou apresentam os seus efeitos apenas tardiamente, o que representa uma dificuldade de que estes problemas, ainda que identificados, sejam associados à intervenção que os causou. Este efeito foi denominado de *iatrogenia*, que literalmente significa “causado pelo que cura”, enquanto o termo *opacidade*

causal denota esta dissociação direta do elemento de intervenção naquele que sofre os efeitos (TALEB, 2012).

A iatrogenia é composta por um problema de agência, ou de risco moral. Ademais, a sua causa imediata está no intervencionismo excessivo, que tem como origem a crença de que a ação humana deliberada deverá sempre contribuir para o funcionamento de um sistema. Por sua vez, esta crença está associada à negação da *antifragilidade* (TALEB, 2012). A antifragilidade é propriedade que se relaciona ao fortalecimento de um elemento ou sistema a partir de choques e fricções, em contraposição aquele que se apresenta como *frágil*, que é danificado por este tipo de agressão. Para fins de contextualização, argumentamos que a *disciplina de proteção de dados pessoais*, e mesmo os sistemas jurídicos nacionais, representam sistemas *frágeis* – pois a fricção com outros elementos e sistemas deverá torná-los menos capazes de atingir os objetivos a que se propõem. Por outro lado, a arquitetura da *internet*, e a informação *digitalizada*, são antifrágeis: a redundância que marca o advento da rede e a natureza da informação gravada em suporte digital representam, diante da sua ubiquidade na vida social, enorme capacidade de permanência e de reposicionamento.

Caso a legislação não seja capaz de tutelar adequadamente atividades de tratamento de dados pessoais, estas não deixarão de existir, pois ocorrem *factualmente* de forma alheia ao direito nacional (exceto, possível e temporariamente, em um estado totalitário, que controle a infraestrutura de telecomunicações). Estas atividades permanecerão, provavelmente ocorrendo de forma mais distante do alcance da legislação, ou ocorrerão de forma a não prover externalidades positivas para a maior parte da sociedade. Por ser *frágil*, o direito que falha em se inserir adequadamente no contexto de um sistema *antifrágil* torna-se obsoleto. Este espaço normativo será preenchido, especialmente diante do forte caráter normativo inerente à arquitetura da *internet* e dos terminais que a ela se conectam. Deparamo-nos com um aparente paradoxo: a falta do direito potencialmente nos conduzirá a um *excesso de regulação* (LESSIG, 1996).

Ao lidar com sistemas de complexidade organizada, *teorias* e *hipóteses*, quando empregadas no desenho de modelos para tomada de decisão, podem representar a fragilização de todo o sistema ao qual se aplicam. Teorias são, por definição, frágeis: apenas existem até que sejam contrapostas suficientemente, quando são então substituídas. No âmbito das ciências sociais, em que os dados passíveis de conversão em conhecimento científico são escassos e inseridos em meio a muito ruído, este problema é potencializado. Assim, um modelo de tomada

de decisão baseado em teorias, se não construído de forma a ser passível de atualização com todo novo conhecimento disponível, frequentemente cristaliza problemas particulares ao aspecto observado, deixando de incorporar informações importantes para o sucesso do objetivo almejado. Por outro lado, a fenomenologia, ou seja, a observação de uma regularidade empírica sem uma teoria prévia visível, pode ser reputada como antifrágil: a observação e sistematização de fenômenos relacionados ao que se deseja aprimorar produz informação robusta, sempre passível de contraposição e adição, que *aprimora* constantemente o conhecimento sobre o sistema e sobre as relações entre as variáveis, tornando o sistema *antifrágil*. Esta abordagem mostra-se adequada também sob o ponto de vista da pesquisa jurídica, pois o direito é um sistema complexo, que se relaciona com outros sistemas complexos; o seu controle democrático depende de uma sistematização *eficiente* do conhecimento acerca dos efeitos de sua atuação, e que seja constantemente adicionada de informação útil e aprimorada.

A fenomenologia, assim, permite abordar problemas observados em sistemas de complexidade organizada de maneira ampla, objetiva e útil. Especialmente, a abordagem fenomenológica nos permite *compreender* aquilo que desejamos ou necessitamos intervir, ao invés de tomá-lo como irracional ou incompreensível. Por outro lado, por não segregar artificialmente o sistema, permite que se adicionem constantemente informações relevantes, que poderão sintetizar sucessivas *teorias enraizadas* (GLASER et al., 1967) – que, ao contrário de hipóteses constituídas *a priori*, se prestam a de fato *explicar* aquilo que observamos, na medida em que nos tornamos progressivamente capazes de apreender o fenômeno. É uma metodologia que não se volta a "ser certa", mas em "acertar". Em outras palavras: deslocamos o foco da variável, para os seus efeitos no sistema – que, ultimamente, é o que entra em nossas vidas, e o que buscamos com todo sistema normativo.

Neste sentido, e no âmbito específico da disciplina de proteção de dados pessoais brasileira, vale mencionar o recente trabalho de Prazeres (2021), que partiu da teoria sistêmica do direito e de teoria ecológica para "destacar a importância de se encarar os problemas jurídicos – e, especialmente, aqueles desdobrados de um ambiente sociocibernético – a partir de uma perspectiva que permita enxergar a sua plena complexidade, ante a qual se revela essencial reconhecer uma dimensão de planejamento" (PRAZERES, 2021, p. 20). Argumenta que o sistema econômico apresentou "extrema habilidade no trato do ambiente digital", ao passo que a implementação "desse modelo exploratório conta com a dificuldade de articulação do direito em torno das expectativas – normativas e cognitivas – que lhe são direcionadas" (PRAZERES, 2021).

Vale dizer, sem a pretensão de aprofundamento do assunto e de tratar da ontologia da *internet*, que a rede nem sempre foi um ambiente dominado pelo mercado, e, em muitos contextos, foi a ele muito mais oposto do que aliado. Por exemplo: as primeiras utilizações da rede eram voltadas a comunicações militares e acadêmicas, sem interesse comercial. Podemos citar, ainda, o caso dos serviços de *streaming* de áudio e vídeo e sua relação com agentes econômicos detentores de direitos autorais: nas décadas de 1990 e 2000, o compartilhamento de arquivos representava uma ameaça para grandes gravadoras, estúdios e produtoras de *software*; hoje, este mercado se acomodou em torno dos serviços de *streaming* e de *software as a service*, e a *internet* representa uma ampliação quantitativa e qualitativa de seu alcance. Utilizando o vocabulário desde capítulo, nota-se que o sistema econômico tem se mostrado *antifrágil* diante das tecnologias da informação, ao contrário do direito, que ainda não foi capaz de inserir o seu programa de sustentabilidade nos diálogos com práticas sociais que se valem de tais tecnologias.

1.2 Narrativas sobre atividades com dados pessoais

O estabelecimento de narrativas que se propõem a explicar um fenômeno – ou, mais notavelmente, que se propõem a transmitir determinado conceitos e *possíveis* explicações para um fenômeno ou conjunto de fenômenos – se constitui como importante elemento da sociedade e da mútua compreensão sobre acontecimentos e como abordá-los. O emprego de metáforas se apresenta como bastante útil neste contexto, e "não são meros embelezamentos estéticos ou sobreposições decorativas a respeito da experiência; elas são parte de nossos sistemas conceituais e afetam a forma com que interpretamos nossas experiências" (LAKOFF; JOHNSON, 1980, pp. 145-6). Na disciplina jurídica de proteção de dados pessoais, o emprego de metáforas é importante e recorrente, possivelmente em decorrência da relativa distância entre os indivíduos amplamente considerados e a lógica de funcionamento por detrás dos sistemas de tecnologia da informação e comunicação que realizam atividades de tratamento de dados pessoais.

1.2.1 Metáforas, pressupostos e objetivos

As metáforas que nos transmitem determinadas ideias carregam o poder de enquadrar nossa maneira de pensar acerca daquilo a que se referem. O pesquisador de filosofia do direito Jack Balkin notou a este respeito que "modelos metafóricos descrevem seletivamente uma situação, e ao fazê-lo ajudam a suprimir conceituações alternativas" (BALKIN, 1998, p. 247). Neste sentido, o emprego de metáforas ao discutir políticas públicas ou outras questões relevantes não pode ser considerado "apenas um empreendimento descritivo, mas também um ato de teorização política com profundas implicações normativas" (SOLOVE, 2006, p. 28), pois o caráter *instrutivo* das metáforas não decorre de seu realismo, mas da forma com que direcionam nosso foco para determinados fenômenos políticos e sociais (SOLOVE, 2006, p. 28), no que observamos o seu potencial de *enquadramento* do processo perceptivo, de articulação argumentativa, e processo decisório. A este respeito, o Guia de Políticas Públicas *ex ante* (Casa Civil da Presidência da República, 2018, p. 53). nota que:

> Uma política é, por definição, fundada em uma compreensão sobre a maneira como os problemas se apresentam e se articulam. Ao priorizar alguns problemas, não se está relegando outros a um segundo plano, mas, sim, organizando-se uma forma de intervir sobre a realidade que pressupõe que alguns desses problemas são causas, enquanto outros, suas consequências.

As narrativas são um poderoso instrumento para a fixação de causas e consequências, e consequente impacto na formulação de políticas públicas e normas de intervenção estatal, pois servem à definição sobre o que será objeto de intervenção e com que objetivos e justificativas.

1.2.3 O Grande Irmão

A literatura supre muitas possibilidades de metáforas. Na disciplina de proteção de dados pessoais e privacidade, e mais amplamente nas discussões sobre relações poder estruturadas na presença de capacidades de vigilância, o *Big Brother* de George Orwell é uma metáfora bastante recorrente na conceituação e tentativa de palpabilização dos problemas que objetiva endereçar. Na distopia totalitária de 1984 (ORWELL, 1949), o Grande Irmão é um governo que sabe de tudo a partir de vigilância constante, e que regula cada aspecto da existência dos indivíduos, requerendo obediência absoluta e abdicação do senso de *privacidade*. Solove (2006, p. 29) observa quanto a este aspecto que "este Estado totalitário aterrorizante alcança o controle mirando na vida privada, pelo emprego de técnicas de poder diversas para obliterar qualquer senso de privacidade", e que "suas técnicas de poder são predominantemente métodos de vigilância". O sociólogo Dennis Wrong (1979, p. 115) coaduna com este ponto de

vista ao notar que "o ápice do horror na anti-utopia imaginária de Orwell é que os homens são privados da própria capacidade de nutrir pensamentos privados e sentimentos opostos ao regime, e ainda mais de agir baseados neles". A teorização no âmbito jurídico e de políticas públicas sobre privacidade e fluxos de informações recorrentemente se utiliza desta metáfora para caracterizar uma noção de *privacidade* e *proteção de dados pessoais* que se relaciona com a autonomia, a partir de um olhar sobre como práticas de vigilância contemporâneas, realizadas por intermédio da infraestrutura da *internet*, teriam o potencial de similarmente afrontar a capacidade de os indivíduos se desenvolverem livremente, com reflexos para a construção social subjetiva e coletiva (SOLOVE, 2006).

Em 1984, há uma combinação de vigilância e força aterrorizante, e os indivíduos não sabem se estão sendo observados – a "teletela" da obra de Orwell atua como o *panóptico*, uma arquitetura de controle idealizada originalmente por Jeremy Bentham (1791) em 1791, e cujo conceito foi exaustivamente debatido no âmbito das ciências sociais a partir especialmente da famigerada obra *Vigiar e Punir*, de Michel Foucault (1977). Arquitetonicamente, o panóptico constitui-se basicamente no projeto de uma prisão em que o pavilhão das celas toma a forma de um círculo ou semicírculo, em cujo centro se encontra uma torre com janelas. Nesta torre está posicionado um único vigilante que é capaz de vigiar a todos aqueles que ocupem as celas do edifício circular, sem, no entanto, ser visto, em razão do efeito da iluminação. Deste fato decorre a noção de que o que importa para os que experimentam esta forma de controle, subjetivamente, não é exatamente a *vigilância*, mas a certeza da possibilidade de estar constantemente vigiado, que exerce um poder normativo – ou de conformação – sobre o sujeito objeto da vigilância. Ademais, o panóptico representa *eficiência* na prática de vigilância também pelo fato de sua arquitetura ser concebida para que um único vigilante, posicionado na torre central, seja capaz de observar o interior de inúmeras celas. Com isso, constitui-se em um aparelho de disciplina, cujo objetivo é assegurar a ordem e a obediência social por meio de sua combinação com a força punitiva.

A *teletela* de 1984 (ORWELL, 1949) representa o meio pelo qual um governo central totalitário, personificado no *Grande Irmão*, atua desta forma sobre os cidadãos: "servindo como uma forma de vigilância de mão única que estrutura o comportamento dos que são observados" (SOLOVE, 2006, p. 31). Contemporaneamente, a adoção da metáfora do Grande Irmão, desta forma, "compreende a privacidade em termos de poder, e enxerga a privacidade como uma dimensão essencial da estrutura política da sociedade. O Grande Irmão tenta dominar a vida privada porque ela é a chave para controlar toda a existência de um indivíduo: seus

pensamentos, ideias e ações" (SOLOVE, 2006, p. 31). Em outras palavras, por estar sempre visível, por estar constantemente sob a realidade de poder estar sendo observado, o indivíduo assimila os efeitos da vigilância em si mesmo, conformando o seu comportamento.

A noção de cultura de vigilância, por outro lado, foi desenvolvida pelo sociólogo William Staples em obra homônima editada em 1997. Staples (1997) argumentou que havíamos criado uma cultura do Grande Irmão, em que todos atuamos como agentes de vigilância e voyerismo. Para acomodar este tipo de concepção, Solove (2006) nota que a metáfora do Grande Irmão passou a ser empregada a partir de uma série de modificações que objetivaram viabilizar a descrição de ameaças à privacidade decorrentes não de um governo central totalitário – um Grande Irmão –, mas de diversos entes privados e públicos que criam e exploram vulnerabilidades relacionadas à informação – como a metáfora dos "Little Brothers" (WHITAKER, 1999; SCHWARTZ, 1999). Neste sentido, o sociólogo David Lyon – que atualmente também emprega e desenvolve o termo *cultura de vigilância*, como veremos adiante – notou que "A visão distópica de Orwell foi dominada pelo estado centralizado. Ele nunca imaginou o quão importante um consumerismo descentralizado poderia se tornar para o controle social" (LYON, 1994, p. 32). Byford (1998, p. 50) argumentou que a "vida no ciberespaço, se deixada desregulamentada, promete apresentar distintos sobretons orwellianos – com a notável diferença de que a ameaça primária à privacidade não vem dos governos, mas do mundo corporativo".

A metáfora em torno do Grande Irmão é útil para a conceituação e compreensão de diversos problemas relacionados à privacidade, inclusive a sua inserção em uma cadeia de relações de poder. No entanto, apresenta notáveis limitações no que se refere aos bancos de dados e fluxos de informação digitalizadas. A este respeito, reproduziremos as palavras de Solove (2006), por sua didática:

> Desenvolvimentos em torno da manutenção de registros não foram orquestrados de acordo com um grande esquema, mas foram amplamente *ad hoc*, surgindo na medida em que a tecnologia interagia com as crescentes demandas do público e das burocracias privadas. Além disso, os objetivos da coleta de dados foram frequentemente benignos – ou ao menos muito menos malignos do que os objetivos do Grande Irmão. Na verdade, informações pessoais têm sido coletadas e armazenadas para uma miríade de propósitos. A história do armazenamento de dados e da produção de banco de dados não é, ao final, uma história acerca do progresso rumo a um mundo governado pelo Grande Irmão ou uma série de Pequenos Irmãos. Ao contrário, é uma história sobre um grupo de diferentes atores, com diferentes propósitos, tentando prosperar em uma sociedade progressivamente baseada em informação. (SOLOVE, 2006, p. 33)

A maior limitação da metáfora do Big Brother, assim, se relaciona à forma de relação de poder que retrata (SOLOVE, 2006). Embora bancos de dados pessoais coletados por dispositivos eletrônicos possam ser utilizados da forma como é retratada por Orwell – por um estado totalitário –, este não constitui o propósito desta estrutura. Ela pode, eventualmente, de forma pontual ou sistemática, ser explorada por um ente desta natureza, ou por outros, para exercer determinadas formas de poder para com as pessoas ou grupos a que se refere a informação. Esta conclusão é consistente com estudos mais recentes acerca do papel da informação enquanto elemento cuja obtenção propicia a criação e a exploração de vulnerabilidades relativamente àquele a que se refere (CALO, 2017).

1.2.4 Kafka encontra Orwell

Notando que a compreensão de nossa sociedade hodierna é um processo contínuo, Solove (2006) propôs a observação de alguns dos fenômenos e problemas relativos a atividades com dados pessoais no contexto do que denomina *dossiês digitais* a partir de outra visão distópica oriunda da literatura: aquela retratada por Franz Kafka em *O Processo* (KAFKA, 1925). A história se inicia quando o protagonista, Joseph K., encontra diversos funcionários públicos em seu apartamento ao acordar, que o informam que ele está preso. Joseph K. fica abismado, e, em certo ponto, questiona: “Eu não consigo me recordar de sequer um delito que me possa ser imputado. Mas, ainda que de menor importância, a verdadeira pergunta é, quem me acusa? Que autoridade está conduzindo estes procedimentos?” (KAFKA, 1925, p. 12). Em meio à sua resposta, o inspetor afirma de forma célebre: “Você está preso, mais do que isso eu não sei” (KAFKA, 1925, p. 12). Os funcionários então simplesmente deixam o apartamento, sem levar K. Durante o resto da história, ele empreende uma frustrada jornada em busca de respostas sobre a situação, como o motivo da prisão e como o caso será resolvido.

Ironicamente, é ele quem toma a iniciativa de buscar o Tribunal responsável pelo processo, que se apresenta como burocrático e misterioso, além de aparentemente ter compilado um dossiê a respeito de Joseph K., que, no entanto, como parte dos registros judiciais, é inacessível ao acusado. Na procura pelo Tribunal, K. encontra advogados, um pintor que trabalha com retratos de magistrados, padres, e diversas outras figuras enigmáticas que, cada qual a seu modo, revelam pequenas peças de conhecimento sobre este obscuro Tribunal e os seus trabalhos. Solove (2006) destaca que apesar de o Tribunal mal ter imposto qualquer

autoridade a K., nem mesmo especificado quando ele deveria comparecer para seu interrogatório, "*ele age como se o Tribunal operasse com regras estritas e empreende todos os esforços para obedecer*" (SOLOVE, 2006, p. 37). O Tribunal aparentemente perde o interesse nele logo após a ocorrência do interrogatório, mas o personagem se torna obcecado com o seu caso, e "*deseja ser reconhecido pelo Tribunal e resolver o seu caso; na verdade, ser ignorado pelo Tribunal se torna um tormento pior do que ser preso*" (SOLOVE, 2006, p. 37).

Em sua busca, K. fica cada vez mais perplexo com a situação: os altos funcionários mantêm-se escondidos; os advogados afirmam ter conexões com funcionários do Tribunal, mas nunca oferecem qualquer comprovação ou resultados; ninguém parece ter contato direto com o Tribunal, cujos procedimentos não apenas eram mantidos em segredo do público em geral, mas também do acusado. K. prossegue de toda forma buscando ser absolvido da imputação de um crime que não lhe foi especificado, por parte uma autoridade que ele não foi capaz de encontrar. Nesta labiríntica burocracia jurídica, K. não consegue obter qualquer progresso: "Progresso sempre foi feito em direção à absolvição, mas a natureza do progresso não pôde nunca ser divulgada. O advogado estava sempre trabalhando na primeira petição, mas nunca chegara a uma conclusão" (KAKFA, 1925, p. 157). No fim, dois funcionários buscam Joseph K. no meio da noite, e o executam.

Solove (2006) argumenta que a metáfora em torno de *O Processo* (KAKFA, 1925) melhor captura o escopo, a natureza e o tipo de relação de poder criada por bases e ecossistemas de dados pessoais. Evidentemente, como *1984* (ORWELL, 1949), trata-se de uma obra de ficção, e, com isso, retrata um cenário em que alguns aspectos da sociedade são destacados, por vezes com recurso a exageros, ao absurdo ou ao satírico. O que é relevante para a sua inserção nos debates ora empreendidos não é um suposto caráter de realidade de seu imaginário, mas o *framework* que este viabiliza para a conceituação de determinados problemas e relações decorrentes dos fenômenos em que dados pessoais estão envolvidos, especialmente em contextos de digitalização e conectividade, e de uma vida social e econômica amplamente dependente (no sentido de *estruturada a partir das capacidades*) destes recursos tecnológicos, dos quais, no entanto, apenas parcela da sociedade conhece e determina a lógica de funcionamento.

O tipo de problema que pode ser conceituado a partir desta metáfora é diverso com relação àquela do Grande Irmão. "Kafka retrata uma burocracia indiferente, em que indivíduos são peões, sem conhecimento do que se passa, e sem voz ou capacidade de exercer controle

significativo sobre o processo" (SOLOVE, 2006, p. 37), o que permite que o julgamento domine completamente a vida de Joseph K.. Com isso, *O Processo* (KAFKA, 1925) "captura o senso de incapacidade, frustração e vulnerabilidade que decorrem do fato de um indivíduo ter grandes quantidades de informações a seu respeito sob o controle de uma grande organização burocrática" (SOLOVE, 2006, p. 38). A todo o momento, algo poderia ser feito contra Joseph K., e ele nem mesmo sabia o quê ou como se defender – seja antes do possível ou ataque, ou após a sua ocorrência. Decisões a seu respeito são tomadas com base nos seus dados – o "dossiê" em posse do Tribunal e ao qual o acusado não tem acesso –, e ele não tem sequer a possibilidade de falar a respeito destas decisões, e, quando finalmente tem, não é ouvido ou não exerce qualquer influência sobre o resultado. Ele está completamente à mercê.

Ao se referir à noção de "burocracia", Solove (2006, p. 38) não se refere a uma determinada instituição, mas a um conjunto de práticas burocráticas, e à forma com que seus processos afetam e influenciam indivíduos que lhe estejam sujeitos; ou seja, às relações, inclusive de poder, que estrutura. Weber notou que a burocracia é "capaz de atingir o grau máximo de eficiência, e sob este aspecto é formalmente o meio mais racional de se exercer autoridade sobre seres humanos" (WEBER, 1978). Por outro lado, que pode se tornar desumanizada ao atuar para eliminar o "amor, o ódio, e todos os elementos puramente pessoais, irracionais e emocionais que escapam ao cálculo" (WEBER, 1978). Neste ponto, é importante observar que burocracias desempenham papeis fundamentais na sociedade moderna, mas, por outro lado, são responsáveis por criar ou exacerbar problemas. Estes aspectos ocorrem sobretudo quando o seu emprego ocorre a partir da redução equivocada de problemas inseridos em sistemas complexos. Por esta razão, este tipo de sistema frequentemente não é capaz de endereçar necessidades de indivíduos em particular – "não pois burocratas são malignos, mas porque precisam agir dentro de limites temporais estritos, recebem treinamento limitado, e frequentemente não são capazes de responder a situações inusitadas de maneiras únicas ou criativas" (SOLOVE, 2006, p. 39).

A responsabilização e *accountability* no âmbito de burocracias são dificultados pela forma com que os processos decisórios são frequentemente mantidos fora do conhecimento do público. Weber (1978) observou que a administração burocrática sempre tende a excluir o público, para esconder tão bem quanto puder o seu conhecimento e a sua ação da crítica. Solove (2006) reflete que

> O problema com bases de dados decorre de sujeitar informações pessoais ao processo burocrático com pouco controle ou limitação inteligente,

> que resulta em nossa falta de participação significativa em decisões sobre nossas informações. Processos de tomada de decisão burocráticos são exercidos cada vez mais frequentemente sobre um crescente âmbito de nossas vidas, e temos pouco poder ou voz dentro deste sistema, que tende a estruturar nossa participação em formas padronizadas que falham em nos permitir atingir nossos objetivos, desejos e necessidades. (SOLOVE, 2006, p. 39)

Neste mesmo sentido, Bauman (LYON e BAUMAN, 2013, p. 235) utiliza-se do termo *adiaphorization*, e especificamente se refere a um cenário contemporâneo de dupla *adiaforização*, que representa a remoção da responsabilidade acerca de processos de categorização (de pessoas) e, ao mesmo tempo, o emprego de um conceito de informação que reproduz o ser humano a este aspecto imediatamente observável, representando ultimamente uma crise de agência, que apenas poderia ser superada a partir de participação dos sujeitos afetados nos processos que lhes afetam.

É este tipo de aspecto que não é capturado na metáfora do Grande Irmão: na distopia de Orwell, o poder depende de uma força e de uma intenção que se unem para dominar e oprimir. O poder, no entanto, não se expressa apenas de formas proibitivas. O poder para Orwell opera como uma força insidiosa empregada para um *desing* específico, enquanto na metáfora de Kakfa o poder é retratado de forma diversa: como não expressando um objetivo, que, se existe, "permanece envolto em mistério" (SOLOVE, 2006, p. 40). Da mesma forma, o poder retratado em *O Processo* não é tão direto e manipulativo quanto na obra de Orwell: o sistema na obra de Kafka difere de nossas noções usuais de estado totalitário. Nas palavras de Daniel Solove (2006, p. 40):

> Joseph K. não foi preso por suas opiniões políticas; nem manifestou o Tribunal qualquer plano para controlar pessoas. Na verdade, Joseph K. estava buscando por alguma razão por que fora preso, uma razão que nunca descobriu. Uma implicação assustadora é a de que não houvesse razão, ou que, se houvesse, esta fosse absurda ou arbitrária.

O que é mais discernível na concepção de Kakfa não é nenhum tipo de objetivo ou motivo nas atitudes do Tribunal e das demais personagens envolvidas na trama, ou em sua forma de exercer poder. O que se destaca na forma de poder retratada em *O Processo* é que ele é menos uma força do que um elemento das relações entre indivíduos e a sociedade e o governo. Estas relações envolvem balanceamentos de poder. *O Processo* ilustra o fato de que o poder não é exercido apenas de formas totalitárias, e que relações com burocracias que sejam desbalanceadas quanto aos poderes envolvidos podem ter efeitos debilitantes sobre indivíduos – independentemente dos objetivos das burocracias (SOLOVE, 2006).

Sob esta ótica, torna-se possível contemplar um aspecto do problema com bancos de dados e ecossistemas de dados digitalizados: eles desempoderam as pessoas, e as tornam vulneráveis por retirar-lhes o poder de tomar parte em contextos em que suas informações pessoais são utilizadas, ou em contextos em que seriam chamados a se manifestar, mas em que suas informações pessoais supriram as necessidades específicas do sistema. Os objetivos dessas "burocracias" são os mais diversos. "[H]á uma teia de decisões irrefletidas realizadas por burocratas de médio escalão, políticas padronizadas, rotinas rígidas, e uma maneira de se referir a indivíduos e suas informações que frequentemente se torna indiferente para com o seu bem-estar" (SOLOVE, 2006, p. 41). Esta *indiferença* também se reflete em infraestruturas de segurança da informação precárias, treinamento insuficiente das pessoas envolvidas em processos que envolvem dados pessoais, tratamento balcanizado de demandas de titulares de dados pessoais, e outros aspectos que culminam em alguns dos problemas enfrentados por titulares de dados contemporaneamente ou os reforçam.

1.2.5 Capitalismo de vigilância

É relevante mencionar a teorização de Shoshana Zuboff sobre o que denominou de *capitalismo de vigilância*. Zuboff (2019, p. 2) emprega o termo *capitalismo* para se referir ao entendimento de que, hodiernamente, a *vigilância* passou a ser empregada como técnica para a acumulação primitiva de dados, especialmente aqueles relativos à experiência humana, que são então de diversas formas transformados e utilizados para finalidades sobretudo econômicas. Zuboff (2019) argumenta que esta lógica tornou-se extrema em virtude do fato de que o valor do conhecimento produzido a partir de dados se relaciona particularmente à sua capacidade *preditiva* quanto a determinado comportamento para determinado público alvo, i.e., um adquirente de dados pessoais, para fins de anunciar um produto para os titulares destes dados, que basicamente deseja que estes anúncios sejam capazes de influenciar pessoas a adquirir o produto; quanto maior for o seu *grau de certeza* quanto à capacidade de estes dados influírem no sucesso da campanha, maior valor atribuirá aos dados. Por outro lado, a agregação de dados aumenta a qualidade e a capacidade preditiva dos dados, o que constitui em um incentivo para que aquelas empresas que lidam com dados pessoais acumulem mais dados, e para que não os compartilhem com potenciais concorrentes. Observa que "*The typical complaint is that privacy is eroded, but that is misleading. In the larger societal pattern, privacy is not eroded but redistributed, as decision rights over privacy are claimed for surveillance capital*" (ZUBOFF,

2019, p. 10). Com isso, ela identifica o problema na *distribuição* do poder decorrente do processamento de dados pessoais.

Este argumento é mais bem esclarecido a partir de um alargamento do exemplo da autora com base na evolução do modelo de negócios do Google. A empresa se constituiu inicialmente com o propósito de fornecer determinados serviços, destacando-se o popular serviço de *busca* na rede. Este serviço constitui, essencialmente, na agregação e organização em um banco de dados, pela empresa, de dados sobre *websites* existentes, além de *scripts*. i.e., instruções para que, a partir de palavras-chave, sejam destacados do banco de dados *websites* que a ela correspondam. Então, uma aplicação, i.e., uma interface, oferece a possibilidade de o usuário inserir uma palavra-chave, desencadeando uma série de operações para que se busque no banco de dados correspondências, que são então apresentadas ao usuário pela interface como *resultados da busca*. Zuboff (2019, p. 4) teorizou que o "*sucesso do Google deriva de uma habilidade de prever o futuro – especificamente o futuro do comportamento humano*". Factualmente, ela observa que, inicialmente, os poucos dados gerados pela interação do usuário com a interface eram vistos como uma espécie de resíduo da operação. Pela própria natureza das redes, algumas informações, como *o quê* um usuário inseriu no campo de busca, e o que ele fez após obter os resultados, i.e., se ele acessou algum dos sites constantes da página de resultados, qual, e quanto tempo levou para tomar esta decisão, são coletados de forma simples pelo sistema. Zuboff (2019, p. 5) observa que o Google empregou estas informações inicialmente com a finalidade de aprimorar os serviços prestados, por exemplo, a partir da compreensão de quais resultados seriam mais relevantes para os usuários, e assim aprimorando os seus algoritmos de busca com base nestas informações. Podemos perceber que isto se refere, ultimamente, à tentativa de aprimoramento da capacidade *preditiva* do algoritmo, i.e., a sua capacidade de apresentar ao usuário aquilo que ele desejaria obter ao acessar o serviço e digitar determinada palavra-chave. Zuboff (2019, p. 5) denomina este momento de "*behavioral value reinvestment cycle*".

A autora argumenta que em um certo momento houve uma guinada rumo ao distanciamento deste ciclo inicial, ocorrendo a descoberta do que ela chama de *behavioral surplus* (ZUBOFF, 2019, p. 5). O *excesso* a que se refere relaciona-se ao fato da descoberta de que estes dados poderiam servir a muito mais propósitos do que meramente oferecer um melhor serviço ao usuário. Em sua concepção mais básica, começou-se a perceber que aquilo que o usuário inseria no campo de busca poderia indicar *mais do que aquilo que ele deseja ao acessar ao serviço*, como, por exemplo, a sua inclinação potencial em adquirir determinado produto ou

serviço. Se, ao adquirir dados de X pessoas, posso ter a expectativa de que 10% irão adquirir o produto, o seu valor pode ser quantificado de forma diversa do que um conjunto aleatório de dados. Em outras palavras: pode ser atribuído a este conjunto um valor que se relaciona a esta *ulterior* utilização. O valor do conjunto de dados se relaciona ao grau de certeza que pode oferecer com relação ao comportamento das pessoas a que se referem, para aqueles que o adquirem com a finalidade de *convencer* pessoas a determinada coisa, e.g., adquirir produto, contratar serviço, votar em um candidato, juntar-se a um grupo extremista etc. Os dados continuam a buscar prever o comportamento do usuário, mas passam a desbordar deliberadamente do contexto da prestação do serviço. Quanto mais específicas forem as predições acerca do comportamento potencial dos sujeitos a que se refere um dado conjunto de dados, e quanto mais confiáveis forem com relação à previsão, mais valiosos são os dados para aqueles que tenham algum interesse em atingir pessoas para influenciar seus comportamentos. Com isso, surge um forte incentivo para a estruturação de sistemas que coletem mais e mais dados: o objetivo deixa de ser prestar determinado serviço; o serviço se torna um mero meio para a obtenção de dados, que podem ser trabalhados para se extrair conhecimento extremamente valioso.

Este processo teria levado a um deslocamento do modelo de negócios ora representado pelo Google, especialmente sob a ótica dos *clientes* do serviço. A empresa passa a buscar obter e prover melhores conjuntos de dados e informações baseadas em dados para os adquirentes destes dados. Assim, se inicialmente os dados pessoais eram “achados naturalmente” (ZUBOFF, 2019, p. 7), constituindo uma espécie de *subproduto* dos serviços oferecidos por APIs, i.e., interfaces, eles tornam-se crescentemente objeto de um esforço consciente de coleta e processamento por parte dos provedores de serviços com o potencial de coletar dados pessoais, especificamente pelo valor gerado pelos dados pessoais coletados durante a interação do usuário com a interface. Ou seja, o serviço que, ao ser utilizado, leva o usuário a fornecer dados pessoais, constitui-se como algo lateral neste fenômeno, o que se torna progressivamente verdadeiro a partir do momento em que determinados serviços deixam de ser inovadores, i.e., podem ser replicados por diversos interessados, com características e valor similar do *serviço* para o usuário. O diferencial passa a ser a quantidade, a qualidade, a velocidade e a variedade dos dados a que um determinado ator tem acesso, que permitem a elaboração de algoritmos com melhor capacidade preditiva, e com isso melhor posicionamento neste mercado.

Acumular dados significa desempenhar bem este papel. A questão que remanesce é: em benefício e em detrimento de quem? A economia movida por dados é um aspecto que decorre

de práticas culturais da contemporaneidade. Agentes de tratamento de dados e titulares de dados são pessoas – cada qual representando diferentes papeis sociais. Por outro lado, nossa sociedade é marcada por desigualdades, em maior ou menor grau. A desigualdade maior neste âmbito, conforme obsesrva Zuboff (2020), é *epistemológica*, ou seja, se relaciona diretamente ao nível de conhecimento de cada uma das partes envolvidas em transações com dados Resta investigar qual é o sentido das categorias jurídicas erguidas pelo GDPR e pela LGPD e que papel eles relegam a estas pessoas no mundo. O poder daqueles que já foram capazes de se estabelecer nesse mercado resta mitigado ou reforçado por esta estrutura jurídica? A teorização de Zuboff contribui sobremaneira para a compreensão dos problemas relacionados a fluxos de informações pessoais na contemporaneidade, especialmente a partir da compreensão de uma relação de poder desigual e a partir de sua conexão com um ou mais modelos de negócios. Por outro lado, as alcunhas empregadas para a descrição deste problema, e o foco em uma relação unilateral e de má-fé, aparentemente o colocam como incompreensível e desumano, quando parece razoável argumentar que se trata de aspectos de diversas práticas sociais que de realizam com diversas finalidades, e que, ainda, podem ser realizadas de diversas outras maneiras, carecendo de uma abordagem que permita a sua compreensão e abordagem conforme os objetivos constitucionais e aspirações da sociedade.

1.2.6 Cultura de vigilância

Em seu texto *Cultura da vigilância: envolvimento, exposição e ética na modernidade digital*, David Lyon (2017) propõe um conceito de *cultura de vigilância*, a que se refere como "expressão guarda-chuva para muitos tipos diferentes de fenômeno, que aponta para a realidade de 'todo um modo de vida' que se relaciona, positiva e negativamente, com a vigilância" (LYON, 2017, p. 162). Assim, diante de uma progressiva mediação digital de nossas relações sociais, os sujeitos são envolvidos, não mais meramente como alvos ou portadores de vigilância, mas como participantes. Descreve a vigilância, assim, como algo não mais apenas extrínseco que se impõe aos sujeitos, mas "algo que os cidadãos comuns aceitam – deliberada e conscientemente ou não –, com que negociam, a que resistem, com que se envolvem e, de maneiras novas, até iniciam e desejam" (LYON, 2017, p. 165). Ainda nesta dinâmica, identifica a aparição de novas questões sobre o envolvimento cotidiano com as mídias digitais, que envolvem aspectos éticos e políticos e que apontam para possibilidades e desafios à cidadania digital. Afinal, "a cidadania digital se conecta especialmente com o que chamam de atos digitais

– jurídicos, performativos e imaginários – e com direito à expressão, ao acesso e à privacidade, além de, atualmente, à abertura e à inovação" (ISIN e RUPPERT apud LYON, 2017, p. 174), e "tanto a vigilância quanto a democracia estão agora mediadas pelo digital" (LYON, 2017, p. 175).

Didático é o contraste que o autor explicita entre *cultura da vigilância* com expressões comumente utilizadas para descrever a fenomenologia relacionada à vigilância. Argumenta Lyon (2017, p. 152) que

> [O] tipo de 'vigilância sem suspeito' executada por agências de inteligência [...] não pode ser compreendido simplesmente nos termos de conceitos mais antigos como Estado de vigilância ou sociedade de vigilância. Agora estes devem ser complementados por um conceito que se concentre mais nos papeis ativos desempenhados pelos sujeitos da vigilância, primeiramente porque tais papeis fazem diferença nos resultados da vigilância.

Um segundo fator é que boa parte daqueles dados é gerada, em primeiro lugar, pelas atividades cotidianas *online* de milhões de cidadãos comuns. Somos "cúmplices", como jamais antes, em nossa própria vigilância ao compartilhar – por vontade própria e conscientemente ou não – nossas informações pessoais no domínio público *online*. *Cultura de vigilância* ajuda a situar isso.

Baseando-se na análise de Charles Taylor (2004; 2007) sobre "imaginários sociais", Lyon (2017) afirma que os imaginários sociais de vigilância têm a ver com entendimentos compartilhados sobre certos aspectos de visibilidade na vida cotidiana e em relações sociais, expectativas e compromissos normativos, ao passo que fornecem uma capacidade de agir, de se envolver e de se legitimar as práticas de vigilância. Essas práticas, por sua vez, ajudam a sustentar imaginários de vigilância e a contribuir para sua reprodução. Por outro lado, "Os imaginários de vigilância oferecem não apenas um sentido do que acontece – a *dinâmica* da vigilância – mas também um sentido de como avaliar e se envolver com ela – os *deveres* de vigilância" (LYON, 2017, p. 161), no que identificamos um potencial *normativo* sobre a conduta dos indivíduos por estes imaginários. Pode-se dizer que terceiros manifestam aprovação ou desaprovação ao tomarem contato com aspectos que com eles são compartilhados, em um processo "autorreflexivo no qual muitos usuários participam e que pode contribuir não apenas para a autoformação individual, mas também para o desenvolvimento de normas e expectativas sociais" (LYON, 2017, p. 168).

As práticas de vigilância, por sua vez, podem ser responsivas, relacionando-se com o ser vigiado, ou iniciatórias, ou seja, quanto a modos de envolvimento com a vigilância. Explorar

a cultura de vigilância atual pelas lentes de imaginários e práticas sociais oferece, dessa forma, maneiras bastante mais abrangentes e compreensivas de avaliar o contexto atual do que conceitos como Estado de vigilância, sociedade de vigilância ou variações da ideia orwelliana de *Big Brother*, como *Big Other* e *Little Others*. Supera, ainda, a dicotomia estabelecida pela dialética do *capitalismo de vigilância*, permitindo vislumbrar categorias menos estanques quanto aos papeis dos sujeitos envolvidos nas práticas de que trata, e ainda alguns dos reguladores (LESSIG, 2006), *affordances* (CALO, 2017) e incentivos de uma maneira geral que para ele se apresentam nesta interação e, mais especificamente, em situações em que cede dados pessoais ou aquiesce com a *datificação* de suas experiências.

Explorar a cultura da vigilância atual sob as lentes de imaginários e práticas forneceria novas formas de pensar a vigilância, a exemplo superando binários conceituais como poder-participação, in/visibilidade e privado-público. O conceito de cultura de vigilância comporta aquele de estratégias de vigilância, a cujo respeito os sujeitos negociam – por exemplo, "percebendo a entrega de dados pessoais como uma troca em benefício próprio" (LYON, 2017, p. 162). Os estudos de Taylor (2004; 2007) demonstram que os "imaginários de vigilância" não são particulares a cada indivíduo, ao passo que são formados pelo sujeito na interação com o meio social, ou seja, de forma discursiva e iterativa. Os imaginários são, portanto, construídos pelo envolvimento cotidiano com a vigilância e suas representações, conforme *percebidas* pelos indivíduos no seio das relações sociais e a partir de suas lentes.

Lyon (2017) argumenta que as instituições incitam diferentes tipos de reação à vigilância, afigurando-se crucial – e "complicado pelo caráter multifacetado das situações em que tal vigilância é experimentada – não reduzir a experiência da vigilância a um formato unidimensional ou binário de 'aquiescência ou resistência'" (LYON, 2017, p. 166). Assim, observa que "hoje, a necessidade de uma ética tão reveladora como normativa é maior do que nunca" (LYON, 2017, p. 171). Com isso, chama atenção para o processo autorreflexivo em que nos inserimos enquanto participantes de uma cultura de vigilância e da ética do "prosseguir" que deriva dessa condição, ou seja, como se comportar nos ambientes virtuais em que se desenvolvem tais práticas.

Lyon (2017), enfim, argumenta que as abordagens atuais, focadas no ponto de vista do *vigilante*, são inadequadas e deixam de se voltar para os seus destinatários – aquelas pessoas *vulneráveis* a partir do uso inadequado de suas informações pessoais na atual *economia movida por dados*. Propõe que, se culturas de vigilância são socialmente construídas, podem ser

desafiadas e reconstruídas. Compreender os desafios éticos e políticos da modernidade digital levaria à inevitabilidade de um conceito como o de cultura de vigilância, que busca abarcar imaginários e práticas de vigilância que produzem complacência, cumplicidade, negociação ou resistência. Assim, essa estratégia, tida por realista, embora centrada na noção de vigilância – e, com isso, de alguma forma limitada pelas relações de poder a ela inerentes – é capaz de apreender outros aspectos que o fenômeno abarca, e, assim, promover a sua compreensão de formas significativas.

1.2.7 O caráter propiciatório

Recentemente, Calo (2017) abordou no âmbito da tutela da privacidade o caráter *propiciatório* de objetos e de sujeitos para com relação à ação de outros sujeitos, i.e., o fato de que um sujeito pode realizar determinadas ações por meio de objetos ou sujeitos, dadas determinadas condições. Esta análise partiu de teoria ecológica da percepção de Gibson (1979), que descreveu a partir do termo *affordance* este caráter que ora chamamos de *propiciatório*. O objetivo do emprego da teoria das *affordances* é viabilizar uma apreensão das potencialidades de objetos ou de sujeitos como meio de ação de um sujeito sobre o ambiente. Um exemplo simples mencionado por Gibson (1979) é a referência a um determinado *plano*, i.e., um terreno, com características de possuir extensão razoável e ser livre de barreiras. Tomando como sujeito uma pessoa com características físicas tais como aptidão plena dos membros inferiores, este plano lhe *propiciará* a realização do ato de correr.

Pontos que merecem nota são: (i) o fato de que a conduta do sujeito não é determinada por aquilo que representa para ele uma *affordance*; o sujeito do exemplo pode, ou não, de fato correr *sobre este plano*, e o fato de este plano com determinadas características estar presente poderá, ou não, sobre ele exercer influências diversas quanto a esta decisão; (ii) um elemento pode *propiciar* diversas ações com relação a um mesmo sujeito, em um mesmo contexto – o que reforça o elemento de não determinismo mencionado anteriormente. No exemplo, e no mesmo contexto, o referido plano também oferece ao sujeito possibilidade de caminhar, e de se sentar com as pernas cruzadas; (iii) no ambiente, frequentemente haverá interação entre diversos elementos, que não apenas propiciam maior âmbito de escolha para o sujeito, como também podem interagir de forma a alterar os *affordances* que ocorreriam na presença isolada de um elemento, e.g., no exemplo acima, caso o plano, embora com extensão razoável, não

fosse livre de barreiras, poderia não ser mais possível correr. Isso não representa necessariamente um impedimento a que o sujeito transpasse a barreira; mas altera as condições, i.e., os custos e os incentivos par que isso seja realizado, e certamente não permite que se corra através da barreira como se esta não existisse. Além disso, conforme mencionado neste ponto, (iv) há uma relação entre *affordances* e *incentivos,* que, porém, não são a mesma coisa: enquanto a um sujeito podem ser possíveis diversas ações, a sua escolha dentre aquelas que são possíveis dependerá de outros fatores, como os seus desejos; os custos para realizar determinado ato; os benefícios potenciais deste ato; o grau de certeza relativamente à realização de perda ou de ganho decorrentes do ato; às demais possibilidades e de como se comparam com a avaliação dos elementos mencionados para o ato considerado; ao grau com que o sujeito se apercebe de quais possibilidades estão presentes; ao grau com que é capaz de compreender as consequências potenciais do ato, i.e., custos, benefícios, certeza de efetivação de resultados.

Ademais, e fundamentalmente, um *affordance* não diz respeito a nenhum dos lados da relação isoladamente: ele descreve aquilo que um elemento propicia a outro elemento, ambos especificados, consistindo, portanto, em característica própria desta *relação*. Neste sentido, vê-se extrema semelhança com a visão empreendida a partir da lógica fenomenológica, consistente com os marcos adotados neste estudo.

Aqui, chamamos a atenção para um aspecto marcante do atual cenário de ameaças a direitos em decorrência da má utilização de dados pessoais: o fato de que os diversos aparelhos eletrônicos e as interfaces digitais por meio das quais interagimos e desempenhamos inúmeras tarefas cotidianas representam meios com características bastante particulares, que representam, contextualmente, *afffordances* diversas relativamente aos sujeitos envolvidos; por exemplo: uma empresa que desenvolva e controle uma determinada assistente pessoal ativada por voz provê algumas *affordances* ao usuário, e.g., acessar determinados serviços. Por outro lado, este aparelho representa *affordances* do controlador da interface operante neste aparelho para com o usuário, ou seja, uma espécie de *capacidade*, ou de *poder*, do controlador para com o usuário.

Vulnerabilidade, ultimamente, se relaciona com poder – de alguém sobre alguém. Por esta razão, Calo (2017) se fiou na teoria das *affordances*, e, especialmente, na menção de Gibson (1979) acerca da possibilidade de que não apenas objetos ou o ambiente representem *affordances* para determinado sujeito, mas que *sujeitos* podem representar *affordances* uns para com os outros – e esta seria a forma mais rica e elaborada de *affordance* disponível no ambiente

(GIBSON, 1979). Estas *affordances* são mediadas por diversos fatores (sociais, físicos, técnicos, culturais etc.), e Calo (2017, p. 601) foca em dois. O primeiro é o papel da informação: uma pessoa apenas representa para outra um *affordance* caso esta possa perceber a primeira como tal – o que remete aos eixos representativos dos significados de informação (HALLINAN e GELLERT, 2020). Gibson (1979) chama de *escondidas* estas *affordances* não identificadas no ambiente. O segundo fator relevante é o papel do direito: um interessado em invadir uma residência pode compreender que ela o *propicia* abrigo, mas o direito de propriedade dirá o contrário (CALO, 2017, p. 602). Ou seja, neste caso, o direito e os remédios jurídicos representam *affordances*, o que permite vislumbrar, ainda, o plexo de influências que atuam sobre o valor efetivo desta *affordance*, considerando a qualidade das leis, a estrutura institucional etc. Neste sentido, a *privacidade* ou o direito à *proteção de dados pessoais* também podem ser concebidas como *affordances*. Assim como outras características ambientais, que *affordances* relacionadas a estes direitos de fato existem varia de acordo com características e capacidades pessoais, e, também, do meio (CALO, 2017, p. 602).

Os problemas aqui descritos se relacionam sob certo aspecto com aqueles do direito consumerista: no cenário atual, apenas empresas que têm como objeto econômico atividades de tratamento de dados possuem *affordances* significativas. Além disso, por terem acesso ao comportamento do usuário e a capacidade de "codificar" o ambiente técnico e jurídico em que transações em meio virtual ocorrem, as empresas são capazes de moldar as *affordances* dos usuários em um grau muito maior do que o reverso. Neste contexto, parece tentador pensar na privacidade e proteção de dados como uma *affordance* capaz de suprir este *gap*, ou seja, capaz de prover equilíbrio a esta relação. Porém, assim como tem sido demonstrado no âmbito do direito do consumidor, as *escolhas* oferecidas pelo direito neste âmbito são frequentemente ilusórias (CALO, 2017, p. 603; CALO, 2013), pois este não é capaz de suprir a assimetria de informação e de poder que está na origem da relação de poder desigual.

Calo (2017, p. 603) defende a utilização da ideia de *affordances* diversos, sob os pontos de vista técnico, jurídico e econômico, para auxiliar na compreensão e estruturação do estudo e da definição de políticas públicas acerca de privacidade, notadamente quando envolve relações de consumo. Neste contexto, e a partir deste *framework*, seria possível questionar se um consumidor ou usuário de determinado serviço é capaz de se aperceber de um conjunto de *affordances* de privacidade que estejam disponíveis, como escolhas de *design* (e.g., criptografia), escolhas de mercado (e.g., produtos que protegem a privacidade ou concorrentes que o façam) e possibilidades jurídicas (e.g., direitos conferidos pela legislação sobre a matéria

e formas de seu exercício), e, ainda, *affordances* que ele representa para os controladores do ambiente virtual. Poder-se-ia questionar, ainda, se estas percepções correspondem à realidade, como estas percepções afetam o comportamento dos usuários e como se manifestam em diferentes circunstâncias para diferentes sujeitos (CALO, 2017, p. 603). Indivíduos amplamente considerados poderiam melhor compreender as potencialidades que determinados aparelhos ou sistemas representam para com relação a si, e melhor avaliar o benefício que obtém de sua utilização com relação a estas potencialidades, e, ainda, como poderiam se envolver de formas mais significativas com práticas que já permeiam a vida social.

1.2.8 Sociedade aberta de dados

Loi et al. (2020) propõem uma abordagem em torno de atividades com dados pessoais em que desenvolvem uma noção diversa de controle e de falta de controle, que se afigura como útil na observação das assimetrias de poder que ocorrem frequentemente neste âmbito. Reconhecem que "a privacidade tem sido compreendida como um resultado de prover indivíduos com mais informação acerca dos usos possíveis de seus dados" (LOI et. al., 2020, p. 13). Por outro lado, que "tem se tornado cada vez mais claro que esta solução não é viável dada a ubiquidade dos dados e transações com dados na economia contemporânea da *internet*" (CATE, 2006), e que decisões relativas a dados pessoais sofrem a influência de incentivos e *nudges* que são incorporados às plataformas digitais (HATRZOG, 2018; WEINMANN et. al., 2016; LOI et al., 2020). A sua proposta se baseia no estabelecimento de cooperativas de dados pessoais em que os membros podem tomar decisões relativas à arquitetura por meio da qual têm experiências virtuais, e granularmente com relação ao uso dos próprios dados. Além disso, os membros da cooperativa teriam iguais possibilidades de gerir conjuntos de dados valorosos para exercer atividades que tragam benefícios para a comunidade.

Loi et al. (2020, p. 14) argumentaram que "um aspecto fundamental de agência no ambiente da internet e na economia do *big data* é deter algum tipo de controle sobre a arquitetura de escolha[1] em que transações com dados ocorrem". Em sua narrativa, a imposição

[1] Esclarece-se que "arquiteturas de escolha online envolvem nudges porque o seu layout visual, a jornada do usuário, as opções padrão, mecanismos de feedback, e painéis do usuário em plataformas da web são elementos de design inescapáveis, e frequentemente constituem nudges, intencional ou não-intencionalmente". Thaler e Sunstein (2008, p. 7) observam que "a choice architect has the responsibility for organizing the context in which people make decisions", e notam que um paralelo crucial entre arquiteturas de escolha e arquiteturas "mais tradicionais" são o fato de que nunca representam um design neutro; ou seja, assim como um arquiteto, ao projetar um prédio, deverá construir algum prédio, o arquiteto de escolhas disporá e organizará as opções de uma

de um modelo de arquitetura de escolha à coletividade – considerando que o número de *arquitetos de escolha* neste âmbito é muito menor do que o de usuários – representa uma espécie de modelo ditatorial, enquanto em um modelo democrático haveria *experts* desenvolvendo alternativas que seriam ultimamente escolhidas por todos os interessados, i.e., usuários e titulares de dados, em uma espécie de votação. Esta solução democrática propiciaria um baixo nível de desigualdade entre usuários e *designers* de serviços movidos por dados, que seria ao mesmo tempo compatível com um grau desejável de eficiência (LOI et al., 2020, p. 14).

Para além de estabelecer aspectos da interface, o modelo também propõe que o controle sobre conjuntos significativos de dados, i.e., acesso a cópia dos dados e autorização para utilizá-los em atividades diversas, seja distribuído entre pessoas da comunidade com aptidão e interesse para tanto (LOI et al., 2020, p. 10)[2]. Partindo do pressuposto de que contemporaneamente dados pessoais representam um importante ativo, e ainda de que a sua inserção em grandes e variados conjuntos é fator preponderante para que deles se possa gerar valor, i.e., realizar empreendimentos que realizem benefícios para a sociedade, uma forma de se distribuir o controle sobre dados que *são gerados* a partir de atividades cotidianas representaria inúmeras possibilidades de impactos positivos para a sociedade sob a ótica da justiça distributiva.

No modelo, a cooperativa como um todo, através de seus membros, define a estrutura e as possibilidades a serem viabilizadas pelo seu ecossistema de compartilhamento de dados, enquanto os indivíduos tomam as próprias decisões relativamente ao compartilhamento dos próprios dados pessoais por meio desta estrutura, granularmente. Desta forma, o objetivo é empoderar os indivíduos com relação ao controle dos próprios dados pessoais, i.e., quais dados podem ser coletados, com quais atores serão compartilhados etc., ao passo que este poder é definido por meio de políticas gerais, códigos de ética e finalidades votados por uma maioria democrática, que, no caso, são os membros da cooperativa de dados (LOI et al., 2020, p. 8). A ideia que embasa a cooperativa de dados pessoais, com isso, se afasta daquela da mera

---

determinada maneira, e, com isso, exercerá um nudge sobre aquele instado a agir no contexto por ele arquitetado. Esta afirmação sobre neutralidade não deve, necessariamente, ser associada à intencionalidade – ao executar determinado projeto, pode-se não ter a intenção de exercer determinada influência sobre aqueles que serão por ele afetados, mas as escolhas de design que necessariamente serão feitas caso se projete algo irão afetar de alguma maneira estas pessoas. A diferença essencial entre um aspecto arquitetônico essencialmente normativo, e um aspecto que constitui um nudge, é o seu grau de irresistibilidade: para que seja apenas um nudge, o elemento deve ser "easy and cheap to avoid" (THALER e SUNSTEIN, 2008, p. 10). Este constitui precisamente o principal limite daquilo que Thaler e Sunstein (2008) chamam de paternalismo libertário.

[2] Loi et al. (2020, p. 10) fazem esta proposição em analogia ao aludido conceito de Rawls denominado property-owning democracy (POD). Rawls advoga por um "regime in which land and capital are widely though not presumably equally held", e "society is not so divided that one fairly small sector controls the preponderance of productive resources" (RAWLS, 1999, p. 247 apud LOI et al., 2020, p. 10).

autogestão da privacidade, que é baseada em "*notions of notice, access and consente regarding the collection, use and disclosure of personal data*" (LOI et al., 2020, p. 9), e não se funda na ideia de que "as pessoas podem decidir por si próprias como sopesar os custos e benefícios da coleta, uso ou divulgação de suas informações" (SOLOVE, 2013) em isolamento de algum tipo de comunidade (LOI et al., 2020, p. 9). Em outro termos, poderíamos dizer que o PDPC constitui uma espécie de enquadramento para o processo decisório do indivíduo relativamente a seus dados, que, conforme o modelo proposto por Loi et al. (2020), seria composto por características e voltado a objetivos determinados de forma razoavelmente afeta aos interesses da maioria dos usuários – ao contrário do que ocorre atualmente, em que tais características são determinadas pela minoria de agentes que controlam a maior parte do ecossistema da economia baseada em dados pessoais.

Em favor de um modelo como o proposto, Loi et al. (2020, p. 11) observam que, atualmente, os usuários de interfaces digitais estão excluídos da camada de interações econômicas em que os lucros são feitos: o uso secundário dos dados. Com isso, usuários podem apenas decidir se consentem com os termos e condições dos ambientes digitais. Poucos agentes são responsáveis pelos usos dos dados pessoais e pela estruturação destes ambientes, diante de alguns limites impostos por jurisdições nacionais (LOI et al., 2020, p. 11). Em contraste com a atual economia movida por dados, em que a autoridade e o controle sobre amplos conjuntos de dados são concentrados em papeis de gerenciamento e propriedade abissais, uma economia em que a maior parte dos dados estejam nas mãos das "cooperativas de dados pessoais" "oferece oportunidades para papeis sociais com um grau intermediário de controle sobre agregados significativos de dados" (LOI et al., 2020, p. 11). Com isso, uma parcela considerável da população exerceria controle sobre este ativo, i.e., dados pessoais que são gerados cotidianamente, endereçando o problema da desigualdade extrema de poder que atualmente se observa quanto a este aspecto.

A justificativa moral para a necessidade de mitigação destas desigualdades se relaciona ao fato de que decisões relativas a arquiteturas de escolha requerem balanceamentos de interesses, e.g., um *design* pode privilegiar a privacidade, enquanto outro pode direcionar os usuários a compartilhar mais dados pessoais. Desta forma, estas escolhas idealmente não devem ser feitas com base somente em aspectos técnicos (LOI et al., 2020; MIRSCH et al., 2017; WEINMANN et al., 2016). Por outro lado, este aspecto é denotativo do poder estruturante das narrativas com relação à in/visibilização de problemas e apresentação de soluções. Loi et al. (2020) argumentam, com isso, que

> The ideia of exercising economic power through not just data choices, but also through data interfaces, realizes one peculiar facet of *predistribution in the data economy*. What gets predistributed, in this case, is not an ordinary economic asset (something with a clear market value), but economic power in a more abstract form: the capacity to influence the behavioral process which generates data and influences their economic value. (LOI et al., 2020, p. 14)

O conceito de predistribuição[3] se relaciona, ainda, a conceituações de modelos distributivos cuja ênfase recai sobre a importância do estabelecimento de mecanismos corretivos do mercado que "atinjam resultados igualitários pela distribuição do controle sobre recursos produtivos, a partir da consideração de que correções *ex-post* constituem alternativas a serem empregadas de forma subsidiária" (LOI et al., 2020, p. 2). Além disso, o modelo proposto se diferencia de outros que propõem uma reorganização da economia de dados sob a lógica da redistribuição da propriedade sobre dados ou das externalidades com eles realizadas isoladamente consideradas, pois também engloba a distribuição desigual do que foi denominado *prerrogativas de autoridade e responsabilidade* na economia movida por dados (LOI et al., 2020, p. 2), através da previsão da concessão de permissões de uso de bases de dados valorosas a membros da sociedade com aptidão para geri-los.

Um dos principais fundamentos da estratégia apontada para lidar com o problema das desigualdades quanto ao controle sobre conjuntos valorosos de dados é a compreensão de que dados pessoais não são bens rivais (LOI et al., 2020, p. 8). Dados digitais podem ser copiados um número indefinido de vezes e armazenados em diferentes locais ao mesmo tempo por

---

[3] A utilização da palavra *predistribuição* decorre do fundamento teórico da proposição de Loi et al. (2020), qual seja, a teoria da justiça de John Rawls. A sua concepção se baseia na *property-owning democracy* (POD) de Rawls e se justifica a partir do segundo princípio de justiça do autor, especialmente a noção de *fair equality of opportunity* (FEO). O ponto de partida é a compreensão de que a possibilidade de um POD em uma economia avançada tecnologicamente é ameaçada pela concentração de poder econômico nesta seara, que é hoje expressa na concentração de poder para moldar os ambientes virtuais relacionados à coleta e reutilização de dados, bem como opções para permitir o seu uso, conforme oferecidas ao cidadão médio (LOI et al., 2020, p. 15). Como solução, propõem a implementação de plataformas cooperativas de dados pessoais (*personal data platform cooperative* – PDPC), e, embasados no referencial teórico adotado, as oferecem como forma de atingir uma ordem institucional análoga ao POD de Rawls. Ressaltam, outrossim, que sua análise não é voltada a acadêmicos das teorias rawlsianas, mas para contribuir para as discussões transdisciplinares hodiernas em torno da política sobre *big data*, e chamam atenção para a inovação central que apresentam: a ideia de (pré-)distribuir os meios produtivos de uma economia movida por dados, deslocando-se da ideia de controle sobre os próprios dados pessoais, para centrar-se naquela relacionada ao controle coletivo sobre a infraestrutura de coleta e utilização de dados pessoais, e dos próprios resultados destes usos, enquanto ao indivíduo torna-se possível, em uma arquitetura de escolha favorável e familiar, o exercício de opções granulares relativamente à utilização de seus próprios dados pessoais (LOI et al., 2020, p. 9).

qualquer indivíduo que detenha o hardware adequado e seja autorizado técnica e legalmente a acessar a informação. Além disso, dados coletados por diferentes agentes ou plataformas frequentemente podem se relacionar ao mesmo indivíduo, ou seja, um titular pode ter a "mesma informação" coletada por diferentes agentes. Disto decorre que dados podem ser compartilhados e agregados de forma a incrementar a sua utilidade econômica e social e sua produtividade (LOI et al., 2020, p. 7; GAL e AVIV, 2020). O fato de um determinado dado pessoal ser coletado por uma empresa que oferece um serviço que o usuário experimenta em ambiente digital, durante a interação com a interface do serviço, não implica no fato de que este dado pessoal não poderia ser empregado para finalidades diversas por quaisquer atores econômicos e sociais que eventualmente tenham acesso a ele, inclusive o próprio titular ou seus mandatários, i.e., o fato de usualmente apenas determinadas empresas utilizarem os muitos dados gerados pela interação cotidiana de pessoas com sensores ambientais e em aparelhos eletrônicos para alguma finalidade, e.g., vendê-los para interessados em anunciar algo a este titular, decorre do fato de esta empresa deter controle sobre a infraestrutura que coleta e processa os dados, e não de um suposto controle sobre os dados pessoais específicos do usuário. Talvez isso nos permita visualizar a razão pela qual propiciar aos titulares controle jurídico, em tese, sobre os próprios dados pessoais, ainda que fosse viável, lhes traria pouco ou nenhum benefício diante da forma com que a economia e cultura contemporâneas se estruturam – o que não equivale a dizer que os titulares de dados não devam poder exercer algum tipo de controle sobre *o quê* é feito com as informações que geram cotidianamente. A questão que se coloca, portanto, é *que tipo de controle* pode conduzir a uma situação de menor grau de desigualdade, e de menos vulnerabilização, do usuário médio de sistema digitais, ao passo que o insere nesta economia contemporânea de maneira significativa, antes de buscar reduzir a sua importância.

Loi et al. (2020, p. 9) observam que uma possibilidade de trazer à tona tais cooperativas, ao menos em parte, existiria na atual legislação europeia (e brasileira) sobre proteção de dados pessoais: se trata do direito à portabilidade, que permite a um titular de dados obter, dentro dos limites estabelecidos, uma cópia de seus dados pessoais em poder de determinado agente de tratamento de dados. Isto permitiria, em tese, que este obtivesse controle sobre esta cópia dos dados – no sentido de determinar os seus ulteriores usos –, que poderia ser levada a uma cooperativa de dados pessoais nos termos delineados. Neste sentido, é interessante observar que a existência de tais cooperativas, inicialmente, não conflita, i.e., pode coexistir, com as atuais empresas e outros atores responsáveis por atividades de tratamento de dados. Ademais, o modelo não pressupõe a existência de uma espécie de "cooperativa central de dados pessoais",

sendo capaz de acomodar, além da coexistência das cooperativas de dados pessoais com outros tipos de pessoas jurídicas, a sua coexistência com outras da mesma natureza ou mesmo outros modelos com objetivos similares, podendo titulares se associarem e se desassociarem livremente conforme seus interesses (LOI et al., 2020).

A polifuncionalidade do direito à portabilidade foi aspecto analisado por Negri et al. (2021), bem como o surgimento de tensões entre pessoa e mercado em decorrência desta polivalência. Foi observado que, se por um lado, ele se volta a promover o controle dos dados por parte daquele a quem os dados se referem, atendendo à perspectiva de proteção à pessoa, por outro também assume o papel de "instrumento de regulação capaz de influenciar a estrutura do mercado digital" (NEGRI et al., 2021, p. 26). O incentivo à circulação dos dados potencialmente promovido pela inserção de um direito à portabilidade poderia ter o potencial de conferir maior autonomia ao direito à proteção de dados pessoais em relação ao direito à privacidade (NEGRI et al., 2021). Alguns dos desafios decorrentes desta tensão entre pessoa e mercado podem ser endereçados a partir da consideração dos mecanismos propostos por Loi et al. (2020), como a distribuição do controle sobre conjuntos de dados e a facilitação da formação de entidades como as cooperativas de dados, o que poderia ser atingido mediante o uso do direito à portabilidade.

Quanto à importância da distribuição do controle sobre conjuntos de dados valorosos, Loi et al. (2020) observam que "While opportunity inequality in the data economy could be ignored in the past, it has become more problematic as the control and exploitation of large data assets is increasingly central to success across multiple social domains" (LOI et al., 2020, p. 12). O controle e a exploração de amplos conjuntos de dados são, com isso, tidos como importantes determinantes de outras oportunidades, i.e., em diferentes áreas da vida social, econômica e política (LOI et al., 2020, p. 12). Como um exemplo extremo, pode-se citar o conhecido escândalo em torno das práticas da empresa *Cambridge Analytica* com relação ao emprego de dados pessoais gerados a partir da interação com redes sociais no contexto político, que denota a potencialidade do controle sobre conjuntos de dados pessoais enquanto importante determinante de acesso a oportunidades no âmbito da política (LOI et al., 2020, p. 7). Em contraste, em uma sociedade em que a maior parte dos dados pessoais sejam geridos por pessoas jurídicas como a do modelo proposto poder-se-ia distribuir a autoridade sobre conjuntos de dados a quaisquer interessados com aptidão para tanto, e que tenham algum interesse sobre o conjunto, i.e., os seus próprios dados pessoais sejam parte do conjunto ou agregado de dados. Seriam criadas condições para a dispersão de oportunidades relevantes entre membros

sociedade, e ao mesmo tempo para a maximização do valor gerado para a sociedade em decorrência do benefício gerado a partir de atividades com dados

Neste ponto, Loi et al. (2020, p. 12) observam que, embora a justificativa normativa principal para a sua proposição seja a sua potencialidade em promover uma relativa igualdade de oportunidades na sociedade, diversos pesquisadores têm enfatizado o potencial das cooperativas de dados em termos de eficiência e para a criação de valor, especialmente em razão de promoverem importantes sinergias entre diversos conjuntos de dados. É amplamente demonstrado pela literatura e por práticas de análises de dados, que a agregação de bases de dados, especialmente de forma sistemática, é capaz de ampliar, quantitativa e qualitativamente, comparativamente ao formato de *silos* em que atualmente são coletados, o conhecimento que deles pode ser extraído. Em pesquisas na área de saúde, por exemplo, tem sido argumentado que cooperativas de dados podem acelerar as pesquisas e suas aplicações clínicas (BLASIMME et al., 2018). Conexões entre dados permitem a obtenção de "super-additive insights" (OECD, 2014). Gal e Aviv (2020) documentaram o potencial das *sinergias* de dados, superior àquele da mera soma dos benefícios isoladamente obtidos a partir das mesmas bases de dados isoladamente consideradas.

Por outro lado, considerando o pressuposto de que a informação, especialmente aquela que foi digitalizada, constitui bem não-rival, plataformas que permitem o sem emprego de formas e para finalidades conforme os interesses dos usuários (ou "titulares"), e sem a pretensão de monopólio que é característica das empresas deste setor, permitem multiplicar práticas valorosas para a sociedade a partir da geração de maior conhecimento e a sua distribuição a mais pessoas capazes de empregá-las para finalidades reputadas relevantes (GAL e AVIV, 2020). Com uma maior aproximação da materialidade que deve ser atingida pelas normas de proteção de dados pode ser possível extrair *affordances* (GIBSON, 1979) significativos. A compreensão da existência de tensões entre pessoa e mercado na origem e polifuncionalidade do direito à portabilidade (NEGRI et al., 2021), associado a proposições como a de Loi et al. (2020), revela alguns destes caminhos. O direito à portabilidade, associado a meios eficazes de controle sobre conjuntos de dados, representa um promissor *affordance* apto a fornecer aos titulares de dados um tipo de controle que estabelece relação estreita e direta com o uso dos dados e com as arquiteturas por meio das quais são coletados e empregados.

Dado este quadro fático, o modelo, ou a utopia proposta por Loi et al. (2020), se afigura como um poderoso norte para um novo olhar sobre a vigente disciplina de proteção de dados

pessoais e outras práticas normativas que possam contribuir para uma melhor posição, jurídica e fática, dos *titulares de dados*, i.e., a maior parte das pessoas, na sociedade, a partir de uma melhor inserção na chamada "economia movida por dados", que se afigura como um aspecto determinante para o acesso a oportunidades de ocupar papeis sociais relevantes. Este tipo de modelo apresenta, ainda, um maior potencial de abordar o tipo de problema de agência (LYON e BAUMAN, 2013, p. 235) identificado no contexto de decisões que impactam pessoas a partir do uso de informações, e enuncia formas diversas de lidar com as externalidades de atividades com dados pessoais. Torna-se possível, ainda, que a concepção de determinados *affordances*, como o próprio direito à proteção de dados pessoais ou o direito à portabilidade, exerça um papel mais significativo para titulares de dados no contexto de estruturas jurídicas que lhes propiciam um espaço decisório e de poder mais amplo.

2 DISCIPLINA JURÍDICA DA INFORMAÇÃO (PESSOAL)

A abordagem fenomenológica deste estudo conduziu a uma revisão bibliográfica acerca da construção do conceito de dado pessoal e conceitos relacionados.

2.1 Informação pessoalmente identificável

Schwartz e Solove (2011) observaram que um dos conceitos mais centrais na regulação acerca de proteção de dados pessoais é o de informação pessoal, e que o escopo destas normas em geral é "ativado" quando informações pessoais estão envolvidas, partindo-se da "assunção elementar quanto à aplicabilidade da norma de que se informações pessoais não estiverem envolvidas, não poderá ocorrer qualquer tipo de dano à privacidade" (SCHWARTZ e SOLOVE, 2011). Pode-se dizer, com isso, que o conceito é fundamental para a definição do âmbito de incidência da norma. Por outro lado, notam que a legislação sobre a matéria, nos EUA e na Europa, não oferece definição para o termo, e que a fenomenologia inerente às atividades com dados pessoais implica em um caráter dinâmico do conceito de informação pessoal. Em razão disso, muitos pesquisadores chegam a sugerir que a dicotomia seja abandonada (OHM, 2010; PURTOVA, 2018). Schwartz e Solove (2011), a seu turno, ofereceram o que chamaram de *conceito 2.0 de informação pessoal*, em que se baseiam em um

standard ao invés de uma regra para acomodar os diferentes contextos em que atividades com dados pessoais ocorrem, para definir a partir deste parâmetro não-binário a aplicabilidade ou não das normas sobre dados pessoais. O conceito por eles desenvolvido é erguido sobre categorias distintas de informações pessoais, que recebem tratamentos jurídicos distintos de acordo com diferentes graus de risco representados por atividades de tratamento de dados pessoais específicas (SCHWARTZ e SOLOVE, 2011, p. 1817).

Em retrospectiva histórica da inserção do conceito de informação identificável de um indivíduo nas disciplinas jurídicas de proteção de dados pessoais e privacidade, Schwartz e Solove (2011, p. 1819) notam que no seminal artigo *The Right to Privacy* escrito por Warren e Brandeis em 1890, os autores "meramente presumiram que a regulação sobre privacidade sempre envolveria informações associáveis a uma pessoa", e, com isso, logicamente "conceberam a privacidade como um direito da 'personalidade'" (SCHWARTZ e SOLOVE, 2011, p. 1819). Em razão de, à época, a "invasão paradigmática à privacidade" se relacionar à publicação de fatos da vida privada de indivíduos pela mídia, ela estaria sempre objetivamente relacionada a uma pessoa, e com isso o ponto não mereceu atenção (SOLOVE e SCHWARTZ, 2011, p. 1819). Devido a este relativamente claro referencial material, o conceito de *informação pessoal*, ou *informação pessoalmente identificável*, foi incorporado ao mundo jurídico sem maiores elaborações, com a presunção de sua referência a objetos concretos e bem delimitados. É relevante observar que esta definição está estreitamente relacionada na sua essência ao *right to privacy* consetuedinário, que fundamentou a sua proteção jurídica.

Este aspecto tornou-se especialmente problemático na década de 1960, com o advento dos computadores, que permitiu que burocracias públicas e companhias privadas processassem dados pessoais de maneiras diversas. Os computadores não apenas aumentaram a quantidade de informação coletada e processada, mas modificaram as formas com que estas informações poderiam ser organizadas, acessadas e procuradas, ou seja, a sua qualidade. Permitiram que a informação passasse a ser organizada, procurada e acessada a partir de múltiplos atributos, ao invés de um único índice, o que transformou a forma com que informações poderiam ser associadas. Neste momento, torna-se evidente que a privacidade não poderia ser tutelada apenas pela proteção e salvaguarda de informações que envolvam o nome ou outras informações diretamente relacionadas a um indivíduo, e o escopo daquelas *informações* que atraem a proteção jurídica declinada aos dados pessoais torna-se significantemente maior e mais controverso (SCHWARTZ e SOLOVE, 2011, p. 1821).

A retrospectiva histórica de Schwartz e Solove (2011) acerca da inserção do conceito de informação nas disciplinas jurídicas relacionadas à tutela da privacidade remonta a outro momento no desenvolvimento legislativo norte-americano sobre a matéria que merece destaque: em 1984, com o advento do *Cable Communications Policy Act*, a ocorrência de informações identificáveis de um indivíduo no âmbito de sua abrangência passam a ser tidas como *gatilhos* para a incidência das regras relativas à tutela da privacidade (SCHWARTZ e SOLOVE, 2011, p. 1824). Se, antes, as normas vigentes naquele país sobre dados pessoais dependiam de contextos de uso de dados pessoais para a sua incidência, no regramento do *Cable Act* o mero fato de determinada empresa coletar determinados dados reputados identificáveis de um indivíduo passa a atrair a incidência das normas específicas sobre proteção de dados pessoais. Schwartz e Solove (2011, p. 1825-6) observam que, se por um lado, a chegada do ano-título da obra de George Orwell chamou grande atenção do público e da mídia para questões afetas à privacidade, por outro, o *tipo de problemas* que se passou a vislumbrar com relação à coleta de informações pessoais sofreu transformações importantes. Na verdade, notou-se que, ainda na década de 1980, as empresas de TV paga perceberam que o serviço que proviam permitia-lhes estabelecer um canal de mão dupla com os seus clientes, que passavam não apenas a receber conteúdo, mas a fornecer determinadas informações, como àquelas relacionadas às suas escolhas de programação, e decorrentes de interação direta com os televisores; com isso, detalhados perfis eram construídos sobre os usuários (SCHWARTZ e SOLOVE, 2011, p. 1826). Com a massificação das tecnologias da informação e comunicação, este tipo de prática, denominado genericamente de "perfilização", tornou-se emblemático do processamento de dados de consumidores para a realização de inferências.

Aparentemente, este elemento levou a que as iniciativas legislativas seguintes se voltassem para a regulação da informação, ao invés das práticas a ela relacionadas, estabelecendo o modelo porvir de regulação sobre privacidade (SCHWARTZ e SOLOVE, 2011, p. 1827). De fato, após o *Cable Act*, o conceito de informação e a ocorrência de práticas de coleta de informações pessoais foram consistentemente empregados como critério de aplicabilidade das normas sobre privacidade. No entanto, esta consistência não se vislumbrou no desenvolvimento dos elementos fundamentais do conceito de informação pessoal ou informação pessoalmente identificável, ou seja, a complexidade deste tipo de informação não foi explorada adequadamente no âmbito jurídico. Se, por um lado, o conceito de informação identificável de um indivíduo for muito estreito, deixará de tutelar aspectos importantes da privacidade, especialmente pelo desenvolvimento da tecnologia; por outro, uma definição

muito ampla pode abranger informação demais, e com isso transformar a legislação sobre proteção de dados pessoais em um sistema jurídico impraticável, e que ao mesmo tempo se impõe a uma infinitude de práticas econômicas e sociais. A legislação sobre privacidade precisa de fronteiras claras, e o conceito de informação identificável de um indivíduo, enquanto central para a atração da aplicabilidade da disciplina, desempenha papel elementar na definição destas barreiras, e, com isso, no erguimento de um regramento jurídico significativo e capaz de tutelar direitos fundamentais conforme as expectativas sociais.

Schwartz e Solove (2011, p. 1829) identificam naquele momento três abordagens predominantes para a definição de conceitos correlatos à noção de informação pessoal, que denominam de abordagem tautológica; abordagem não-pública; e abordagem de tipos específicos. As duas primeiras se apresentam como *standards*, enquanto a última é construída sob a estrutura de uma regra jurídica. De forma simplificada, podemos considerar que um *standard*, ou um princípio, se refere a uma norma jurídica de realização do comando no maior grau faticamente possível, enquanto uma regra enuncia uma norma do tipo *tudo ou nada*. A primeira abordagem identificada, denominada tautológica, define informação pessoal como qualquer informação que identifique um indivíduo. A sua virtude está em sua natureza relativamente aberta, o que decorre de seu caráter principiológico, que representa capacidade de desenvolvimento e flexibilidade em tese capazes de acomodar mudanças tecnológicas que alterem a materialidade regulada pela disciplina de proteção de dados pessoais. O seu problema está na ausência de definição da informação pessoal ou explicação de como poderia ser identificada (i.e., reconhecida em suas representações), podendo-se dizer que, "em seu núcleo, esta abordagem simplesmente afirma que informações pessoalmente identificáveis são informações pessoalmente identificáveis" (SCHWARTZ e SOLOVE, 2011, p. 1829), e desta forma é pouco útil na distinção de informações pessoais de informações não pessoais.

A segunda abordagem, denominada *não-pública*, define como pessoal aquela informação que não seja pública, realizando a distinção, portanto, de forma negativa. Esta abordagem é problemática, especialmente, por sua incapacidade em, de fato, referir-se a aspectos relativos ao caráter de *pessoalmente identificável* da informação, pois centra-se em outro aspecto, qual seja, o da publicidade e acessibilidade ao público da informação. Desta forma, já neste nível elementar mostra-se incapaz de atuar sobre os problemas que se apresentam contemporaneamente com relação ao uso de dados pessoais, em que a quantidade e qualidade dos problemas é determinada em grande medida por aspectos como a *agregação*

de dados e inferências realizadas a partir de dados primários, frequentemente tidos por inócuos em si (SOLOVE, 2006).

A terceira abordagem principal identificada por Schwartz e Solove (2011, p. 1831) é a dos *tipos específicos*. Esta abordagem, diferentemente das duas anteriores, se estrutura como uma *regra* jurídica, ou seja, especifica situações em que estarão presentes informações pessoais a partir de limites definidos *ex ante*. Desta forma, o primeiro problema que se relaciona a esta abordagem é em seu potencial restritivo na definição de informação pessoal. Este problema é exacerbado, e provavelmente intransponível, no âmbito de problemas relativos a informações pessoais no contexto dos computadores e da *internet*: por exemplo, conforme observa Calo (2017), o que é determinante para a importância da tutela jurídica quanto ao uso de informações pessoais é a identificação das situações em que a pessoa a que se refere a informação puder ser vulnerabilizada por aquele que a obtiver; por outro lado, a agregação de dados (SOLOVE, 2006) e características da informação em suportes digitais (VILLARONGA et al., 2018) são determinantes para a ocorrência destas situações, e a sua compreensão fundamental para a identificação destas ocorrências, uma vez superado o "paradigma do segredo" (SOLOVE, 2006).

Diante deste quadro, há pesquisadores que questionam a pertinência da distinção, no âmbito jurídico, entre informações pessoais e informações não-pessoais[4], o que, em conjunto com considerações acerca da importância de dados de outras naturezas (em princípio, não pessoais), como os "agrodados"[5], demonstra a pertinência de que esta questão seja enfrentada. Quando da elaboração da LGPD, estes conceitos não foram discutidos – sempre foi considerada uma definição *tautológica*, em que dado pessoal é a informação referente a uma pessoa, recaindo as discussões sobre a abertura ou não desta definição. Também conforme veremos adiante, há uma distinção fundamental entre tratar do conceito de dado pessoal, enquanto categoria de informação, e tratar diretamente da conceituação de *informação pessoal*, ou *informação pessoalmente identificável*. Em outras palavras, embora muita atenção tenha sido dada na última década às discussões em torno da proteção jurídica em torno de *dados pessoais,* ao objeto desta regulação, que, conforme extraímos das discussões empreendidas por Schwartz e Solove (2011), vem sendo delimitado deste a década de 1980 a partir do conceito de

[4] CF. OHM, 2010; PURTOVA, 2018.
[5] CF. CARBONELL, 2016.

informação pessoalmente identificável, muito pouco foi-se capaz de construir acerca do conceito de *informação* que subjaz estes demais conceitos.

Schwartz e Solove (2011, p. 1873) observam que as normas sobre proteção de dados pessoais se estruturaram, a partir deste momento inicial, em basicamente dois eixos: uma regulação *reducionista* nos Estados Unidos, e uma regulação *expansionista* na União Europeia. É interessante notar que, no momento em que os pesquisadores fizeram estas observações, o GDPR, que *expandiu* o escopo da regulação sobre dados pessoais da União Europeia, ainda não estava em vigor. O reducionismo a que se refere no tipo de normativa existente nos EUA está na usual consideração de que informações pessoais são exclusivamente aquelas que se referem a um indivíduo imediata e atualmente identificável; pela frequente adoção do critério de tipos específicos; e pela regulação setorial. Por outro lado, a abordagem expansionista adotada na União Europeia define "dado pessoal" como aquela informação que se refere a uma pessoa natural identificada ou identificável, e abrange quaisquer atividades que envolvam dados pessoais, que atraem a incidência da norma. É notável que esta definição foi repetida pelo GDPR, e também pela Lei Geral de Proteção de Dados Pessoais brasileira de 2018, o que demonstra a pertinência e a contemporaneidade do debate.

Embora reconheçam que a abordagem europeia está em maior sintonia com o atual estágio tecnológico, a principal crítica de Schwartz e Solove (2011, p. 1875) à abordagem expansionista europeia é ao seu idêntico tratamento jurídico-normativo de toda informação que se refira a uma pessoa natural identificada ou identificável, independentemente, por exemplo, de a pessoa ser *identificada* ou *identificável*; dos esforços e custos associados à possível identificação; dos riscos e benefícios inerentes à identificação; ou do contexto em que ocorre a atividade de tratamento. A origem deste idêntico tratamento foi identificada no *Bundesdatenschutzgesetz*, ou Lei Federal de Proteção de Dados Alemã de 1977, que definiu "informação pessoal" indistintamente como informação relativa tanto a indivíduos identificados, quanto identificáveis (SOLOVE e SCHWARTZ, 2011, p. 1874). Desta forma, enquanto no mundo fático observa-se um *continuum* de informações identificáveis que incluem diferentes formas de informações pseudônimas e anônimas, com diferentes graus de esforço associados à identificação da informação a uma pessoa, e variados graus de risco associados a esta possível identificação, dentre inúmeras outras nuances, a adoção da abordagem expansionista na identificação de informações pessoais para aplicação indistinta das regras jurídicas inerentes é uma abordagem manifestamente equivocada (SCHWARTZ e SOLOVE, 2011, p. 1876). Mais recentemente, esta abertura semântica do conceito de dado ou informação

pessoal revelou novos desafios diante da função de definição da materialidade de incidência da norma que também incumbe ao conceito, para além de definir o seu âmbito de incidência (HALLINAN e GELLERT, 2020).

Como um exemplo extremo das dificuldades inerentes, e para ir além da distinção básica entre "grandes" e "pequenos" agentes de tratamento de dados, cite-se o caso em que um agente processe dados relacionados a uma pessoa *identificável*: o adjetivo implica em uma situação de fato em que a pessoa *ainda não foi identificada*, ou seja, não *está associada à informação* neste momento. Uma disciplina jurídica de proteção de dados pessoais que trate este caso indistintamente daquele em que a informação se refere a uma pessoa identificada – como é o caso do GDPR e da LGPD – exigirá que o titular destes dados (i.e., a pessoa *identificável* a partir da informação em questão) tenha direitos como ser informado da ocorrência da atividade de tratamento, acesso aos dados e correção dos dados. Estes direitos deverão ser viabilizados pelo agente de tratamento de dados, e o deverão ser tão somente com relação ao titular do dado (i.e., apenas a pessoa a que se refere a informação identificável deverá tomar conhecimento e poder acessar e corrigir as informações). Para tanto, o agente de tratamento deverá *identificar* o sujeito que antes era apenas *identificável*, o que, dentre outros efeitos, incrementa o risco a que o titular de dados está sujeito, e também o risco do passivo que os dados representam para o agente de tratamento de dados. Este exemplo de ocorrência indesejável e contraproducente diretamente decorrente do tratamento jurídico-normativo indistinto de todo tipo de "informação pessoal" parece tornar auto evidente a sua impropriedade e a importância do atingimento de outros contornos regulatórios.

2.2 O abrangente conceito de dado pessoal

A literatura jurídica em torno da conceituação de dado e informação pessoal ou pessoalmente identificável gravita, frequentemente, em torno do elemento da *identificabilidade* de um indivíduo a partir de um determinado dado, e dos esforços e medidas tecnológicas necessários para a identificação ou reidentificação de pessoas a partir de bancos de dados. De fato, conforme observa Purtova (2018, p. 42), pode-se dizer que há hoje relativo consenso quanto à impossibilidade fática de que dados reputados pessoais sejam transformados em anônimos de forma absoluta, i.e., de forma a não mais poderem ser utilizados para identificar o indivíduo que originalmente identificavam antes do processo de anonimização. Alguns dos

pesquisadores que observaram este aspecto são Ohm (2010), Sweeney (2000) e Schwartz e Solove (2011), enquanto Tene e Polonetsky (2013) notaram como análises de *big data* tornam insignificante a distinção binária entre informação identificável ou não-identificável de um indivíduo. Por outro lado, o que estes autores parecem deixar de lado é que o "problema com o conceito de informação pessoal vai além da simples identificabilidade, porque o segundo elemento essencial do conceito de 'dado pessoal', i.e., a relação da informação para com uma pessoa, também é problemático" (PURTOVA, 2018, p. 42).

Purtova (2018), assim, pretende contribuir para o tema a partir do argumento de que "na era da Internet das Coisas, datificação, análises avançadas de dados, e tomada de decisões com base em dados, toda e qualquer informação se refere a uma pessoa no sentido da disciplina jurídica europeia de proteção de dados" (PURTOVA, 2018, p. 42). O seu argumento não é em prol de um conceito restrito de dado pessoal, tampouco por uma restrição do escopo de abrangência da proteção de dados. Na verdade, partindo do pressuposto de que todos os dados apresentam o potencial de causar danos a pessoas – o que ela utiliza como critério para a definição de *pessoal* –, todo dado deveria desencadear algum tipo de proteção contra possíveis impactos negativos. Por outro lado, se para a pesquisadora o problema central não está na abrangência extensiva da norma, ele se encontra no fato de, diante da ampla aplicabilidade da norma, impor-se de maneira generalizada um regime de direitos e obrigações intenso e não-escalável, que resulta em poucas possibilidades de que o GDPR seja implementado de maneira significativa, e pode ainda representar a ocorrência de efeitos negativos, especialmente à medida que a norma se torne progressivamente abrangente pela expansão material qualitativa e quantitativa da ocorrência de atividades sob sua jurisdição, conforme o critério de aplicabilidade estabelecido.

Na legislação europeia sobre proteção de dados pessoais, o conceito de "dado pessoal" é central para a definição de seu escopo de incidência, pois apenas quando estão em jogo "dados pessoais" aplicam-se as suas obrigações e os seus direitos previstos. A definição do GDPR é ampla, flexível e adaptável ao contexto tecnológico hodierno. As referências a "pessoa natural identificável" e a "informação que se refere a uma pessoa natural" "convidam a interpretações quanto ao que constitui uma possibilidade relevante de identificação, e o que constitui uma relação relevante entre a informação e um indivíduo" (PURTOVA, 2018, p. 44). De fato, o Recital 26 do GDPR adota um teste de razoabilidade com relação à probabilidade de identificação, seja pelo controlador, seja por um terceiro. Este teste leva em consideração, sobretudo, o estado tecnológico atual enquanto viabilizador ou não desta identificação

(PURTOVA, 2018, p. 44). Desta forma, aparentemente se está a falar de um juízo acerca de que tipo e grau de relação da informação com um indivíduo se afiguram como significativos, bem como determinar se esta relação está presente em circunstâncias específicas, pelo que se pode dizer que estes elementos convidam a uma interpretação contextual do sentido de "dado pessoal" (PURTOVA, 2018, p. 44). A partir especialmente de uma contraposição das definições de "dado pessoal" e "dado anonimizado"/"anonimização" da LGPD, Bioni (2018) notou que pressupõe-se um critério de proporcionalidade a delimitar o âmbito de incidência da norma quanto ao critério de "identificabilidade". Desta forma, o caráter contexto-dependente da delimitação do termo tem a consequência de alçar a *status* dinâmico a qualificação *pessoal* do dado, e, com isso, um mesmo conjunto de dados pode, inicialmente, não reunir as características que o qualifiquem como pessoal, porém, sob a perspectiva de outrem ou na medida em que novas tecnologias de processamento de dados sejam desenvolvidas ou que estes dados sejam com outros combinados, *tornar-se*, ou *aparentar ter sido durante todo o tempo*, um dado "pessoal" (PURTOVA, 2018, p. 47).

Assim, para além da dicotomia pessoa identificada/identificável, é fundamental notar outros aspectos da definição legal do termo, e observar como atuam contextualmente. Tanto o GDPR, quanto a LGPD, ao definir "dado pessoal", referem-se a *qualquer* informação que se refira a uma pessoa natural. Não há, com isso, distinção entre a natureza, conteúdo ou formato da informação. Quanto à natureza, podemos afirmar que as normas não distinguem, por exemplo, entre informação falsa e verdadeira, ou entre objetiva e subjetiva (inclusive opiniões e avaliações) (PURTOVA, 2018, p. 48). Além disso, o *conteúdo* da informação não é relevante para a sua qualificação como pessoal: ela pode ou não conter informações sobre a vida familiar ou profissional do indivíduo, ou ainda sobre qualquer outra coisa, desde que *se refira a pessoa natural*. Quanto ao formato, ou ao suporte da informação, também não há ressalvas: a informação pode ser "alfabética, numérica, gráfica, fotográfica ou acústica" (PURTOVA, 2018, p. 49); pode ser mantida em papel, memória de computador, meios biológicos ou quaisquer outros; estruturada ou não estruturada etc. Para além dessa abertura, o GDPR e a LGPD utilizam o termo "informação" como auto-evidente em seu significado, e não esclarecem como "informação" se relaciona a outros termos por vezes intercambiáveis, como "dado", "significado", "conhecimento" ou artefatos contêineres de informação (CDs, livros etc.) (PURTOVA, 2018, p. 49).

A partir destas observações, Purtova (2018, p. 50) nota que se torna viável argumentar que, ultimamente, *tudo* constitui, ou ao menos contém, informação. Este conceito, não apenas

no campo jurídico, varia ao longo do tempo e conforme o tratamento que lhe é conferido por diferentes disciplinas, como a filosofia, a psicologia e a cibernética, conduzindo ao que Burgin (2010) chamou de "perplexidade dos estudos da informação", e conforme se discute mais detidamente na seção seguinte a partir do estudo de Hallinan e Gellert (2020). Essa variação do conceito não prejudica – antes parece favorecer – este argumento pela expansividade extrema do conceito.

Purtova (2018, p. 51) nota que, embora no âmbito da disciplina jurídica de proteção de dados o conceito de informação como *dado acrescido de significado* tenha sido amplamente adotado, muitos estão em desacordo com este enfoque e rejeitam a ideia de que a informação seja sempre significativa para aqueles que a utilizam. Neste sentido, Hildebrandt (2018) estabelece uma comparação ao afirmar que todo tipo de organismo, ainda que não se comunique da forma com que seres humanos são capazes, dependem de informação para sobreviver e prosperar; igualmente, sistemas baseados em inteligência artificial igualmente dependem do processamento de informação para existirem, funcionarem e se aprimorarem. Nem no caso de organismos biológicos, nem no caso de máquinas inteligentes, a informação implica necessariamente na atribuição de significado. Basta pensar em uma colônia de formigas: através de mecanismos biológicos, os espécimes difundem informação que é captada por seus pares na forma de estímulos, por exemplo, para escapar de perigo iminente. Sistemas de inteligência artificial, a seu turno, usualmente associam pontos de informação em decorrência de alguma similaridade ou distinção entre elas, não se afigurando como um elemento fundamental a atribuição de significado.

Este debate se insere naquele que busca determinar se a informação é algo que pertence exclusivamente ao domínio humano, ou se pode ser encontrado em qualquer parte do universo (BURGIN, 2010). Purtova (2018, p. 51) observa que cada vez mais físicos definem o mundo físico como composto de informação, e alguns chegam a argumentar que "informação estrutural e cinética é um componente intrínseco do universo [...] independentemente de alguma forma de inteligência poder ou não o perceber" (BURGIN, 2018). Diante deste contexto, Purtova (2018 p. 53) argumenta que "não podemos mais dizer que alguns dados não têm significado, pois perdemos para os computadores o monopólio de decidir a este respeito", e, com isso, seria melhor assumir que "tudo são dados, e todos os dados têm significado; assim, tudo é ou contém informação" (PURTOVA, 2018, p. 53). Sob o ponto de vista material, esta afirmação pode ser associada a práticas que implementam técnicas de mineração e exploração de *big data*, como

documentado, por exemplo, por Zarsky (2016), que são capazes de deduzir informação inteligível para seres humanos a partir de dados (informação?) que antes não o eram.

O vocábulo "relativo a" que conecta *informação* a pessoa natural também se afigura como aberto no regime jurídico do GDPR e da LGPD. Por outro lado, de acordo com o Working Party 29 do GDPR, o significado deste vocábulo é "crucial, uma vez que é muto importante descobrir quais são as relações/links [entre a informação e o indivíduo] que importam e como distingui-los" (PURTOVA, 2018, p. 53). Se em algumas situações é relativamente óbvia esta relação, em outras não se pode afirmar que seja auto-evidente. Um contexto em que este último cenário é frequente é aquele em que a informação se refere imediatamente a um objeto, i.e., ao valor de uma casa, ou informação relativa a um processo ou um evento em que há intervenção humana. Nesses casos, a relação da informação com uma pessoa pode ser dita indireta, pois decorre da relação ou da interação do ser humano com estes objetos ou outros elementos. O Working Party 29, aparentemente, considera suficiente uma relação indireta do indivíduo com a informação para que esta seja reputada pessoal, e, com isso, atraia a incidência do GDPR (PURTOVA, 2018, p. 54).

Purtova (2018) nota que a informação pode se referir a uma pessoa a partir de diferentes condições alternativas, a saber: de seu conteúdo, de seu propósito, ou de seu resultado, ou seja, informação "referente a" implica em um conjunto de ocorrências mais amplo do que "sobre" uma pessoa, que no primeiro está contido. Além disso, como os três possíveis aspectos de correlação com uma pessoa não são cumulativos, mas alternativos, as possibilidades de que uma informação se refira a uma pessoa são ampliados. Importante notar, ainda, que uma informação pode se referir a uma pessoa, no sentido legal, sem que seja sobre esta pessoa, o que se verifica nos casos em que a informação é utilizada com o propósito de subsidiar uma decisão automatizada que impactará um ser humano, ou os seus interesses e complexo de direitos. Em síntese, pode-se dizer, com isso, que nem toda informação se refere a um indivíduo em razão de seu conteúdo. No ponto em que Purtova (2018) reforça a importância de um conceito de *data-driven agency* (HILDEBRANDT, 2016), argumentando que, contemporaneamente, "qualquer informação pode se referir a uma pessoa em razão de seu propósito, e toda informação se refere a uma pessoa em razão de seu impacto" (PURTOVA 2018, p. 54), ou de seu impacto sobre o complexo de direitos e interesses de uma pessoa, ainda que de forma diminuta, i.e., um tratamento diferenciado (PURTOVA, 2018, p. 56). Esse tipo de conceituação direciona para uma compreensão de dado pessoal a partir de um *standard*,

conforme proposto por Solove e Schwartz (2011), e ainda uma compreensão de matizes de vulnerabilidade em ocorrências e transações que envolvem dados pessoais (CALO, 2017).

Diante deste cenário, dois problemas se colocam: por um lado, é preciso que se compreenda como atividades de tratamento de dados impactam as pessoas individual e coletivamente consideradas. Então, torna-se necessário distinguir entre aqueles que geram impactos inaceitáveis, e quais são aceitáveis. Em ambos os casos, será necessário levar em consideração que benefícios decorrem destas atividades. Por fim, é preciso que sejamos capazes de distinguir quais são os destinatários destas externalidades positivas e negativas – que poderão, inclusive, ser distintos, o que gera um problema de agência, ou de risco moral. Pode ser que esta estruturação do impacto das atividades com informação revele a desnecessidade da qualificação de pessoal para a amplitude de casos prescrita pelo GDPR, e sugira caminhos relacionados à uma melhor justiça distributiva com relação a possibilidades de exercer controle e desempenhar atividades produtivas a partir de conjuntos de dados. Por outro lado, coloca-se o problema de definir um tipo de regime jurídico de proteção de dados que seja escalável, o que implica na necessidade de diferentes regimes jurídicos para diferentes atividades, com diferentes impactos. Sobretudo diante do intenso e custoso regime imposto pelo GDPR, esta falta de gradação, associada à grande abrangência da norma, podem conduzir a uma sobrecarga que dificulte ou impossibilite o atingimento de resultados significativos para a sociedade a partir daquele regime jurídico e dos que adotam os mesmos princípios.

Neste ponto, para Purtova (2018, p. 76), o maior problema identificado a partir de sua análise não é exatamente o amplo escopo da disciplina jurídica de proteção de dados pessoais devido ao amplo e aberto conceito de dado pessoal; o problema jaz na alta intensidade do *compliance* requerido pelo GDPR, indistintamente para quase todo tipo de agente de tratamento de dados. Vale dizer que, quanto a este aspecto, a LGPD adotou postura idêntica, embora contenha autorização legal para que a Autoridade Nacional de Proteção de Dados Pessoais regulamente condições distintas, sobre aspectos específicos de suas obrigações, para alguns tipos de agentes de tratamento de dados, como aqueles com atividade de pequena relevância. No entanto, há importantes razões para se argumentar que este escopo de regulação não é suficiente, e que as nuances que envolvem atividades com dados pessoais são diversas (SCHWARTZ e SOLOVE, 2011; GAL e AVIV, 2020).

Com relação ao peso da adequação do GDPR, Purtova (2018) rememora que o Regulamento se fia em uma noção de "cultura de *compliance* contínuo", o que implica, por

exemplo, em observância ao princípio de *accountability*, que pressupõe permanente capacidade de demonstrar a adequação à norma. A maior parte das obrigações impostas aos agentes de tratamento de dados não é meramente de se abster, mas são obrigações de fazer frequentemente custosas e de difícil realização (GAL e AVIV, 2020). Considerando que o GDPR e a LGPD se aplicam a praticamente qualquer pessoa que realize atividade com dados pessoais, a qualquer tempo e para qualquer finalidade, com a ampliação do escopo de incidência da norma, pode-se vislumbrar um cenário em que suas disposições – e sérias sanções – sejam passíveis de aplicação ampla. Em um cenário extremo, a adequação à norma seria muito difícil ou impossível, levando alguns controladores a, ao invés de realizar análises significativas de impacto, adequarem-se às suas disposições de forma dissimulada ou balcanizada (SOLOVE, 2006), e outros à sua mera desconsideração (PURTOVA, 2018). Ultimamente, conduziria a um contexto de descrédito e decadência de sua capacidade normativa, com a fragilização dos direitos que visa a tutelar. Nas palavras de Julie Cohen (2018), conforme rememoradas por Purtova (2018, p. 77), a disciplina de proteção de dados se tornará uma espécie de "teatro de Kabuki", que distrai tanto usuários e titulares de dados, quanto reguladores, daquilo que realmente está acontecendo com relação ao uso de dados contemporaneamente.

2.3 O conceito de informação

Recentemente, seguindo este programa de pesquisa, Hallinan e Gellert (2020) conduziram estudo inédito em que foram capazes de apreender problemas até então invisíveis no GDPR, a partir de análise específica do *conceito de informação* na disciplina jurídica europeia de proteção de dados pessoais. Foi identificado que, no GDPR – e presumidamente em outras disciplinas jurídicas de proteção de dados pessoais que foram erigidas à semelhança de seus princípios fundantes e relevantes sob os aspectos ora evidenciados – , o conceito de informação desempenha dois papeis distintos: por um lado, enquanto um critério de aplicabilidade da disciplina; por outro, como objeto de regulação da norma. Além disso, foi possível observar que os limites substantivos do conceito de informação diferem consideravelmente com relação a cada um destes papeis, o que têm efeitos importantes relativamente à eficácia da disciplina, especialmente o potencial de torná-la inapta a tutelar liberdades fundamentais na contemporaneidade diante de sua inadequação, i.e., incapacidade de endereçar os problemas para os quais foi elaborada e sobre a materialidade sobre a qual deve atuar.

O fato de o conceito de informação ser central na disciplina de proteção de dados não deve causar espanto; afinal, é a substância, a coleta, a troca e a manipulação da informação que provê a *ratio* para a existência de normas sobre proteção de dados (HALLINAN e GELLERT, 2020, p. 271). De fato, seja no GDPR, seja na LGPD, o conceito de informação é integrante-chave da conceituação de *dado pessoal*, o que significa que o conceito tem papel importante na definição do objeto das normas. Porém, o GDPR não traz qualquer definição de "informação" (HALLINAN e GELLERT, 2020, p. 271), como não o faz a LGPD. Jurisprudência e academia também não proveem conceituação consistente sob o ponto de vista multidisciplinar (HALLINAN e GELLERT, 2020, p. 272). Como observamos especialmente a partir das discussões empreendidas por Schwartz e Solove (2011), Warren e Brandeis (1989) não precisavam se preocupar com o mesmo tipo de problema relacionado à utilização de informações pessoais com que lidamos contemporaneamente, e desta forma o sentido do conceito também não se afigurara problemático. Por outro lado, o direito alemão se incumbiu de assimilar pessoa *identificada* à pessoa *identificável*, e, nos EUA, na década de 1980, o mero fato de lidar com dados pessoais passou a atrair a incidência de disciplinas jurídicas específicas (SCHWARTZ e SOLOVE, 2011). Contemporaneamente, diante de características muito diversas quanto ao uso de informações pessoais – inclusive quanto ao que constitui "informação pessoal" –, o conjunto deste desenvolvimento normativo parece ter culminado em uma situação que não pode ser superada a partir dos mesmos pressupostos que foram e seguem sendo adotados nas legislações sobre proteção de dados pessoais, que o foram considerando o tipo de problema que se enfrentou quando da associação da tutela da informação pessoal ao *right to privacy*.

Analisando a questão sob o ponto de vista do regulamento geral europeu sobre proteção de dados pessoais (GDPR), Hallinan e Gellert (2020) consideraram que esta falta de conceituação expressa de um aspecto central da disciplina, que ao mesmo tempo se afigura como multifacetado e complexo, poderia representar graves problemas para a eficácia e aplicabilidade de norma. Por esta razão, dedicaram-se a esmiuçar o emprego do conceito de informação no GDPR, e, com isso, propuseram e demonstraram a pertinência de três teses cumulativas: (1) o fato de que o conceito de informação desempenha dois papeis diversos na disciplina, a saber: o de critério de aplicabilidade, e o de objeto de regulação; (2) que as fronteiras substantivas dos dois conceitos diferem em cada um destes papeis, i.e., são dois conceitos de informação diversos, que se relacionam a fenômenos substantivos diversos; e (3) que as diferenças substantivas entre estes dois conceitos de informação são significativos para

a eficácia do GDPR enquanto instrumento jurídico voltado a regulamentar fluxos de informação (HALLINAN e GELLERT, 2020, p. 272).

Enquanto um critério de aplicabilidade, o conceito de informação funciona para definir se o GDPR, ou a LGPD, se aplicam *rationae materiae* (HALLINAN e GELLERT, 2020, p. 273). Considerações sobre este primeiro papel do conceito de informação foram mais comuns nos debates acadêmicos e de políticas públicas, em comparação ao segundo papel que será tratado adiante, o que se supõe devido ao fato de constar expressamente de definições legais de ambos os diplomas. No caso da LGPD, destaca-se a sua aparição no conceito de dado pessoal: "informação relacionada a pessoa natural identificada ou identificável" (artigo 5°, inciso I). Por sua vez, lê-se do artigo 1° da LGPD que "Esta Lei dispõe sobre o tratamento de dados pessoais, inclusive nos meios digitais" (BRASIL, 2018), e no *caput* do artigo 3° lê-se que "Esta Lei aplica-se a qualquer operação de tratamento" (BRASIL, 2018), desde que realizada no território nacional ou relativamente a titulares localizados no Brasil, ou oferte bens ou serviços no território nacional. No GDPR, o Art. 4(1) define dado pessoal como "qualquer informação relacionada a uma pessoa natural identificada ou identificável" (UNIÃO EUROPEIA, 2016), enquanto o Art. 2(1) dispõe que "Este regulamento se aplica ao tratamento de dados pessoais realizado em parte ou totalmente por meio de meios automatizados" (UNIAO EUROPEIA, 2016). Vemos com isso que o conceito de *informação* é chave para a definição da aplicabilidade das normas, pois compõe o conceito de *dado pessoal*, que, por sua vez, é objeto cuja natureza atrai a aplicabilidade da norma. De fato, conforme observam Hallinan e Gellert (2020, p. 274), "a presença ou a ausência de informação determina que matérias podem, ou não podem, ser dados pessoais, e aos quais o GDPR e suas previsões substantivas possam se aplicar".

O segundo papel identificado do conceito de informação é implícito no GDPR e na LGPD. Enquanto *"objeto de regulação*, o conceito de informação funciona como uma matéria em torno da qual os princípios substantivos do GDPR foram desenhados, e em relação aos quais estes princípios atuam" (HALLINAN e GELLERT, 2020, p. 274). Embora não tenha sido explicitado nos diplomas, "este conceito deverá ter tomado alguma forma na mente do legislador para que este tenha se engajado na escolha e no desenho das previsões substantivas" (HALLINAN e GELLERT, 2020, p. 274). De fato, razão assiste às pesquisadoras em sua observação neste aspecto: pelo papel central que ocupa a noção de *informação* para o planejamento e estruturação da norma, que visa precisamente a tutelar informações qualificadas como "dado pessoal" por referirem-se a pessoa natural identificada ou identificável, um conceito de *o quê* se estava a endereçar deverá ter sido considerado para que se concebesse e

estabelecesse mecanismos jurídicos aptos a exercer controle sobre este *algo*, sobre este *tipo de objeto* que se está a buscar exercer controle com vistas a proteger direitos fundamentais. Um exemplo didático sobre este ponto é dado por Hallinan e Gellert (2020, p. 274): o GDPR e a LGPD concebem e concedem aos titulares de dados um direito de *acesso* a seus dados pessoais – isso significa que o titular tem o direito de solicitar ao, e obter do, agente de tratamento de dados uma cópia de seus dados pessoais sob tratamento por este agente. Para prover esta cópia, o agente de tratamento de dados precisa realizar uma série de procedimentos relacionados à matéria, ou ao substrato, da informação; em outras palavras, deverá agir sobre o objeto *informação* relacionada à pessoa natural identificada ou identificável. Desta forma, a concepção de normas jurídicas voltadas a regular este tipo de procedimento deverá ter levado em consideração algum tipo de *conceito de informação* para que se pudesse conceber que tipo de mecanismos poderiam interagir com este objeto, e que gerariam interações passíveis de proteger os direitos intendidos. O mesmo poderia ser dito sobre um direito de informação sobre atividades de tratamento de dados pessoais, e de exclusão de dados pessoais nas hipóteses em que é cabível.

Resumidamente, distingue-se duas *funções* do conceito de informação na disciplina de proteção de dados pessoais estruturante do GDPR e da LGPD: enquanto critério de aplicabilidade, o conceito exerce uma função normativa ao definir se determinado substrato se qualifica para receber a proteção declinada pela norma; enquanto objeto de regulação, o conceito exerce uma função descritiva quanto a um substrato com caraterísticas específicas em torno das quais a normativa foi erigida, e sobre a qual deverá atuar (HALLINAN e GELLERT, 2020, p. 275).

Uma observação superficial levaria à conclusão de que o conceito de informação relativamente a estes dois papeis convergiria quanto a seus aspectos substantivos, i.e., o conceito teria os mesmos limites substantivos (fáticos). No entanto, Hallinan e Gellert (2020) observaram que, em muitos aspectos, estes conceitos diferem quando são observados os fenômenos concernentes a cada um dos papeis do conceito de informação, i.e., os conceitos apresentam limites substantivos diversos na definição de *informação* e se relacionam, com isso, a diferentes fenômenos substantivos. Para viabilizar a análise deste conceito conforme empregado no GDPR, Hallinan e Gellert (2020, p. 276) elaboraram uma moldura teórica a partir da fenomenologia da informação. Partiram do pressuposto básico de que, no nível mais alto de abstração, a "informação é um recurso para resolver a incerteza". Esta definição é consistente com o emprego do conceito no GDPR e na LGPD: de fato, quando se fala que uma informação

que identifica ou pode identificar uma pessoa natural é um dado pessoal, aparentemente se está a falar de uma informação que permite resolver uma incerteza, objetivamente referindo-se a um indivíduo específico. A moldura teórica proposta identifica três diferentes eixos de compreensão do conceito de informação e delimitação daquilo que está nele abrangido, o que viabiliza a compreensão dos limites materiais e características dos dois papeis distinguidos.

No primeiro eixo, o aspecto central para a definição de se algo é, ou não, *informação,* se refere ao "grau com que a informação deve semanticamente relacionar-se com significado presente no mundo" (HALLINAN e GELLERT, 2020, p. 276). A partir das lentes deste eixo, podemos observar que nem todos os campos do conhecimento conceituam informação como algo que necessariamente transmita significado imediato acerca do mundo, como é o caso de conceitos matemáticos de informação com enfoque nas relações probabilísticas entre sistemas independentemente de conteúdo semântico (HALLINAN e GELLERT, 2020, p. 276). Ademais, a "informação pode diferir no grau de estruturação necessário para transmitir conhecimento a um agente. A informação pode ser deliberadamente estruturada para transmitir significado fluidamente – como em uma frase factual – ou a informação pode ser menos estruturada, requerendo a adição de mais, ou menos, *frameworks* interpretativos complexos para a extração de significado" (HALLINAN e GELLERT, 2020, p. 277).

O segundo eixo de avaliação se relaciona ao "grau com que a informação deve ser armazenada em, ou transferida entre, meios, i.e., mídias, suportes específicos" (HALLINAN e GELLERT, 2020, p. 277). Certas conceituações de informação apresentam enfoque na necessidade de que a informação esteja armazenada em determinado suporte para ser considerado como tal. Por exemplo, algumas definições no campo da ciência da computação podem ter como critério o fato de o objeto esteja armazenado em determinado meio físico criado pelo homem, i.e., *hardware* específico, para que seja considerado "informação" neste campo do conhecimento. Nas ciências biológicas, por outro lado, o conceito de informação pode se referir a algo que ocorra naturalmente no ambiente, como por exemplo *informação* contida em tecidos de animais ou plantas. Hallinan e Gellert (2020, p. 277) mencionam como exemplo os "anéis" que podem ser observados em troncos de árvores, que existem independentemente de intervenção humana e que estão correlacionados com a idade da planta – e, com isso, podem, *sob algum tipo de conceito de informação*, ser considerados informação a este respeito, independentemente da capacidade humana de apreendê-la.

Por fim, o terceiro eixo permite observar a inserção, na conceituação de informação, de critérios relacionados ao "grau com que a informação deve se relacionar à cognição humana". Estes critérios envolvem o estabelecimento de parâmetros de avaliação do objeto com enfoque na cognição, capacidade criativa e percepção humanas. A título de contextualização, podemos recordar-nos de um ultrapassado, porém tradicional conceito jurídico: o de "homem-médio". A noção de como um "homem-médio" ter-se-ia comportado diante do acesso a determinado objeto (uma informação em potencial) que lhe teria possibilitado não cometer um ato ilícito, por exemplo, é crucial para o veredito de seu caso, e ultimamente envolve a definição de se aquilo a que ele teve acesso consubstanciou-se, ou não, em informação (compreensível e útil para o indivíduo – no que enxergamos o eixo de compreensão delineado). Por outro lado, conceitos no âmbito da biologia que levam em conta a *informação* contida no DNA – por exemplo, acerca da herança genética e tradução genotípica e fenotípica – considerarão informação aquilo que está no DNA, e que opera independentemente da cognição humana. No exemplo relativo aos "anéis" em troncos de árvore, sob este último eixo, a qualificação como informação poderia depender da capacidade humana em apreender conhecimento a partir deste elemento.

A partir destes eixos, de fato, uma comparação entre os dois conceitos de informação revela múltiplas diferenças. Resumidamente, Hallinan e Gellert (2020, p. 304) distinguem dois tipos-chave de distinções: (1) diferenças nos graus de flexibilidade do conceito; e (2) diferenças nos limites substantivos do conceito. Com relação às diferenças entre os conceitos em termos de flexibilidade, foi observado que uma comparação entre os contextos de cada um permite inferir um grau extremo de flexibilidade do conceito de informação quando empregado no papel de critério de aplicabilidade, em comparação com o seu emprego enquanto objeto de regulação. Isto decorre do fato de que este conceito, eventualmente, determina a aplicabilidade da proteção jurídica declinada pela disciplina de proteção de dados pessoais, e *contextualmente* se torna relevante sempre que riscos a direitos individuais na sociedade da informação são identificáveis (HALLINAN e GELLERT, 2020, p. 305), de forma ampliativa pela posição de vulnerabilidade atribuída ao *titular* (CALO, 2017). Sob esta ótica, a flexibilidade do conceito de informação enquanto critério de aplicabilidade é extrema. No caso do conceito de informação empregado no papel de objeto de regulação também há algum grau de flexibilidade; porém, os conceitos e relações estabelecidos em referência a este papel têm limites definidos no direito e na linguagem natural que não podem ser ignorados. Por exemplo, para que o agente de tratamento de dados realize determinados deveres legais em relação ao titular, ou para que o titular exerça direitos

sobre estes dados, é necessário que se leve em consideração a capacidade humana de associação da informação ao indivíduo, o que, sob as lentes do terceiro eixo acima referido, denota uma definição restritiva.

Com relação às fronteiras substantivas específica dos dois conceitos, são identificadas variações em cada um dos três eixos voltados à análise conceitual da informação. Relativamente ao primeiro eixo, enquanto a informação como critério de aplicabilidade engloba todo tipo de informação semântica, o conceito de informação como objeto de regulação se refere tão somente àquelas informações semânticas que se consubstanciam em fatos sociais, i.e., que constituam "dados pessoais" passíveis de que se atue sobre eles. Esta observação decorre naturalmente da necessidade de que se identifique os "dados pessoais" sobre os quais agentes de tratamento, titulares ou reguladores devem atuar. Considerando o segundo eixo, observa-se que a informação como critério de aplicabilidade se refere apenas a meios que propiciam a cópia e a transferência de informação (meios artificiais), i.e., uma informação que não esteja fixada e seja objeto de atividade de tratamento não pode atrair a incidência da norma. Relativamente ao terceiro eixo, o conceito de informação como critério de aplicabilidade requer envolvimento humano para a definição do contexto de seu processamento, enquanto como um objeto de regulação requer, adicionalmente, uma habilidade humana cognitiva capaz de perceber e compreender a informação (HALLINAN e GELLERT, 2020, p. 306), para sobre ela atuar de maneira significativa, inclusive para o exercício de direitos por parte do titular ou atuação da atividade regulatória.

Como resultado dessas diferenças quanto ao grau de flexibilidade e às fronteiras dos conceitos de informação nestes dois papeis distintos, é previsível que o vazio substantivo entre ambos se tornará mais acentuado com o tempo. Por um lado, o escopo do conceito de informação como um critério de aplicabilidade deverá se expandir, o que decorre do fato de que as relações sociais hodiernamente têm se tornado progressivamente mediadas por informação, sobretudo comunicada por meio de interfaces de dispositivos eletrônicos. Com isso, mais situações com o potencial de representar ameaças a direitos fundamentais em decorrência de possibilidades de criação e exploração de vulnerabilidades (CALO, 2017) deverão ocorrer, desta forma atraindo a aplicação da disciplina de proteção de dados pessoais, considerando a abrangência dos três eixos delineadores do conceito de informação (HALLINAN e GELLERT, 2020). O conceito de informação que desempenha este papel, com isso, deverá se expandir caso se intenda que a norma prossiga sendo capaz de tutelar indivíduos em situação de vulnerabilidade. Por outro lado, o conceito de informação enquanto objeto de regulação deverá

permanecer relativamente estático, o que se infere a partir das limitações em termos de flexibilidade dos princípios substantivos da disciplina de proteção de dados pessoais (HALLINAN e GELLERT, 2020, p. 307), ou seja, da facticidade e materialidade em que visa interferir. Enquanto alguns dos aspectos problemáticos com o conceito de dado pessoal, como a uniformidade da incidência das normas de *compliance* contínuo (PURTOVA, 2018), sejam em alguma medida aptas a serem endereçadas por meio do exercício da competência regulatória das Autoridades em matéria de proteção de dados pessoais, aqueles ora identificados requerem intervenção legislativa devido ao caráter naturalmente expansivo do seu âmbito de incidência legalmente referenciado e do âmbito de atuação materialmente delimitado, o que, no momento, não está no horizonte do legislador (HALLINAN e GELLERT, 2020, p. 317).

Por causa destas diferenças, formula-se a hipótese de que o GDPR, assim como a LGPD, aplicar-se-ão a tipos de informação para os quais os seus princípios substantivos não foram desenvolvidos, ao passo que deixarão de exercer influência sobre diversas situações em que a vulnerabilidade do titular de dados de fato está presente, embora seja atraída a sua incidência a partir do conceito de informação enquanto critério de aplicabilidade ante a existência de relação de vulnerabilidade, devido à inaptidão para atuar sobre a informação considerada sob o aspecto de objeto de incidência das disposições normativas. Desta forma, cabe verificar como estas diferenças substantivas entre os dois conceitos podem representar ameaças à eficácia da norma. Estas diferenças, assim, não constituem "curiosidades acadêmicas": a divergência é um fator causal de uma série de problemas concretos relativamente à eficácia da disciplina como um instrumento jurídico voltado a tutelar direitos no âmbito dos fluxos de informação (HALLINAN e GELLERT, 2020, p. 318).

2.4 Vulnerabilidade e privacidade

A partir das discussões empreendidas, parece possível argumentar que a oportunidade de vulnerabilização é o elo entre a informação e o sujeito que, na relação de fundo, justifica a caracterização como dado pessoal que atua no papel de indutor da incidência da norma. Por outro lado, a atribuição do caráter de vulnerabilidade a partir de categorias jurídicas estanques, i.e., a partir da própria característica de titular de dados, não vai além da definição tautológica (SCHWARTZ e SOLOVE, 2011), e, com isso, não traz possibilidades de melhor definição

deste âmbito de incidência. Parece que adequado seria a adoção de uma noção de vulnerabilidade que não se revelasse por meio de rótulos, mas de camadas (LUNA, 2009).

Zuboff (2020) argumenta serem aspectos centrais dessa situação de vulnerabilidade o déficit epistemológico dos *titulares*, e o que chama de lógica de acumulação de dados caracterizadora das atividades econômicas de empresas que exploram atividades relacionadas a dados pessoais. Lyon (2017) supera algumas dicotomias comuns no tratamento do tema ao escrever sobre os fenômenos que subjazem a economia movida por dados e a forma como as pessoas, para além de *usuários* ou *titulares*, são tidos como *sujeitos* que tomam parte em uma *cultura*, cuja centralidade, sob esta ótica, jaz na própria interação destes sujeitos com as estruturas, ou *imaginários sociais*, que caracterizam nosso momento histórico, de intensificação do processo de virtualização por meio da digitalização. Neste ponto, vale dizer que a *virtualização* é processo muito anterior à ideia de digitalização da informação: uma pintura rupestre, por exemplo, é *virtualização* de homens de outrora, que deixaram ali a sua marca, sem, no entanto, permanecerem presentes fisicamente. A *digitalização* da informação e dos espaços de interação social, que se relaciona de múltiplas formas com a virtualização, é processo distinto, com características particulares.

Embora ainda se centre marcadamente na ideia de *vigilância*, quando parece claro que a fenomenologia inerente aos fluxos informacionais apresenta muitos outros aspectos, Lyon (2017), aparentemente, propõe uma saída para aquilo que Solove (2020) identificou como *o mito do paradoxo da privacidade*. O *paradoxo da privacidade* corresponde ao argumento de que as pessoas em geral afirmam valorizar a sua privacidade, porém as suas escolhas diuturnas retratariam o oposto. A aceitação deste paradoxo conduz a teorias com pressupostos reducionistas, como aquelas centradas na vigilância ou em suposta unilateralidade das relações entre *titulares* e *agentes de tratamento* de dados pessoais, quando a observação das práticas sociais em que se inserem as atividades com dados pessoais denotam uma complexidade que se perde na binaridade imposta por estas categorias jurídicas, como sugere a própria obra de Lyon sobre a *cultura* que envolve práticas e imaginários sociais de "vigilância".

Indivíduos são a todo o momento confrontados com situações em que tomam decisões, e as tomam em situações específicas, tendo em vista objetivos e riscos específicos, com acesso a diferentes graus de informação e dados estímulos diversos. Em outros contextos, as atividades de coleta de dados pessoais ocorrem de forma alheia aos indivíduos, e a sua utilização acontece de forma *diferida*, ou seja, em momento e/ou em contexto diverso daquele em que ocorreu a

coleta. Assim, a economia movida por dados e os aspectos cuja regulação é necessária para endereçar as assimetrias ora discutidas são muito mais do que a eventual cessão de dados pessoais para finalidades diversas. Muitas atividades de coleta de dados sequer dependem da interação consciente do indivíduo com algum objeto ou pessoa. Conforme notam Villaronga, Kieseberg e Li (2018), "infelizmente, nossas atuais legislações não são hábeis para manejar as complexidades e desafios da inteligência artificial", e "uma área em que nossa legislação atual é insuficiente é na regulação sobre privacidade". Tudo isso reforça a necessidade de nos centrarmos nas externalidades das atividades de tratamento de dados, para compreender como ela se relaciona com os diversos atores sociais e como podemos erguer nomas jurídicas adequadas a este contexto, ou seja, que deem conta destes múltiplos fenômenos. O foco da regulação sobre tratamento de dados, seja para a proteção da privacidade, seja para a promoção de outros valores fundamentais, depende de outro tipo de compreensão e enquadramento dos problemas resultantes de tais atividades de tratamento e as práticas econômicas e sociais que lhe são correlatas.

Calo (2017) propõe uma distinção analítica entre *tornar-se* vulnerável, i.e., tornar alguém vulnerável, e *explorar* uma vulnerabilidade, seja esta natural ou induzida. Descreve, ainda, a relação entre privacidade e vulnerabilidade como um ciclo vicioso *ou* virtuoso: quanto mais vulnerável uma pessoa é, menos privacidade tende a gozar, enquanto a falta de privacidade deixa o indivíduo exposto a maior vulnerabilidade e exploração. A recíproca é verdadeira.

Ao afirmarmos que alguém está vulnerável, queremos dizer que esta pessoa está vulnerável a sofrer algum tipo de dano. No âmbito da proteção de dados pessoais, como em outras disciplinas jurídicas, concebe-se a vulnerabilidade como decorrente de um determinado *status* ou relação específica (CALO, 2017, p. 592). Esta concepção geralmente se fia em relações estanque, e.g., no direito do consumidor, o *consumidor* é reputado *vulnerável* em decorrência de determinada posição contratual; na disciplina de proteção de dados pessoais o *titular* é considerado vulnerável em decorrência de dados a que se lhe referem serem utilizados ou guardados por outrem etc. Esta concepção binária, i.e., da vulnerabilidade como presente ou ausente em determinados contextos ou relações não é, porém, a única forma de enxergá-la: podemos concebê-la de com mais nuances, de forma a levar em conta as circunstâncias em que se insere. Neste sentido, Calo, (2017, p. 593) observa que a literatura jurídica já foi capaz de absorver este tipo de noção de vulnerabilidade, referindo-se ao trabalho de Luna (2009), para quem a vulnerabilidade deve ser compreendida não a partir da ideia de que alguém *é* vulnerável, mas, sim, a partir da compreensão de situações específicas que *tornam* alguém vulnerável, ou

seja, uma vulnerabilidade experimentada em *camadas* e expressa *contextualmente*. A apreensão desta noção não-binária da vulnerabilidade é elementar e permite enxergar o problema a partir de novas lentes. Calo (2017, p. 593) destaca dois pontos que decorrem desta compreensão: o primeiro é que ninguém é completamente *invulnerável* em nenhum momento: somos todos vulneráveis em graus diversos, de acordo com circunstâncias particulares. O segundo é que a vulnerabilidade não é completamente um produto do acaso: as circunstâncias que se relacionam com a vulnerabilidade, i.e., que a instalam, que permitem que seja explorada ou que a afastam, podem ser controladas ou manipuladas, e, assim, a própria vulnerabilidade também o pode (CALO, 2017, p. 594).

Ao analisar os diversos papeis da *privacidade*, Calo (2017) compreende que ela pode ter função dúplice, ou seja, ser um *escudo* ou uma *espada* com relação à própria vulnerabilidade ou de terceiros. A função de escudo pode ser compreendida a partir da noção de que obter informação sobre um indivíduo representa um poder potencial sobre aquele indivíduo, ou seja, uma vulnerabilidade potencialmente explorável. Calo (2017, p. 594) utiliza como exemplo a história do personagem Super-Homem, em que o personagem Lex Luthor foi capaz de enfrentar em pé de igualdade o super-herói pelo conhecimento de sua fraqueza. Neste caso, a informação foi utilizada para perpetrar um dano físico, como poderia ser feito caso se soubesse sobre a grave alergia de um indivíduo a amendoins. O conhecimento pode ser utilizado, ainda, para fins de *chantagem*: o Dr. Martin Luther King, Jr. foi vítima de escutas ilegais pelo FBI e, em seguida, vítima de chantagens, sob ameaça de divulgação de supostas provas de relações extraconjugais. De fato, Ohm (2010) especula que sobre praticamente qualquer indivíduo deverá haver um fato cujo compartilhamento amplo lhe seria ruinoso, o que a torna extremamente vulnerável a quem seja eventualmente capaz de obter esta informação e comprová-la. O terceiro exemplo mencionado por Calo (2017, p. 595) da utilização da informação sobre indivíduos para explorar vulnerabilidades é no contexto da persuasão. Este é um dos aspectos mais explorados no âmbito da *economia movida por dados*, em que estudos baseados em economia comportamental buscam prever os comportamentos *irracionais* a que estão mais propensos determinados indivíduos, a julgar por informações que se obteve sobre este indivíduo ou grupos a que pertence, com o intuito, por exemplo, de fazê-lo adquirir um produto ou serviço de forma impulsiva.

Em todos estes contextos, a privacidade poderia agir como um escudo, representando uma espécie de barreira que dificulta a instalação, a descoberta ou a exploração de vulnerabilidades. Assim, poder-se-ia caracterizar a privacidade, em um nível básico, como uma

affirdance cuja função seria viabilizar a minimização da exploração e da instalação da vulnerabilidade pelo escondimento da própria vulnerabilidade (e.g., da informação sobre a localização ou alergia a amendoins) ou pela proteção da informação que, se descoberta, poderia nos tornar vulneráveis no momento (CALO, 2017, p. 596). Há situações em que esta situação é subvertida: a privacidade pode agir para esconder a criação e a exploração de uma vulnerabilidade. Um exemplo foi a utilização, pelo Google, de normas do GDPR para reduzir a competição e incrementar as barreiras de entrada em mercados em que atua (GAL e AVIV, 2020).

Calo (2017, p. 599) argumenta que a privacidade pode também obscurecer o conceito de vulnerabilidade. Haveria um custo à imposição da etiqueta "privacidade" a contextos em que o dano decorre, na verdade, de uma situação de vulnerabilidade. De fato, como se argumenta ao longo deste trabalho, aparentemente as categorias jurídicas de *titular* e *agente de tratamento* de dados, no mais das vezes, apenas enunciam uma determinada situação fática, e determinam um plexo de direitos que se pretende a ela inerente. No entanto, pode ser que os conceitos que atraem a incidência da norma sejam mais amplos do que aqueles que caracterizam objetos passíveis de intervenção, seja no caso do GDPR (HALLINAN e GELLERT, 2020), seja no caso da LGPD. Por outro lado, o tipo de vulnerabilidade a que determinado titular de dados está sujeito pode ter o uso de sua informação como um elemento lateral, ou seja, que é empregado com o condão de vulnerabilizá-lo, mas que não representa o objeto que se desejaria regular. Por exemplo, Calo (2013) argumenta que o tipo de manipulação de mercado hoje empregada é muito mais sofisticada do que a de outrora, e que empresas estudam fastidiosamente os clientes, sendo então capazes de não apenas fazer uso de uma compreensão genérica de limitações cognitivas para influenciá-los, mas de identificar elementos específicos capazes de atiçar um determinado sujeito, e mesmo de *buscar este cliente* no momento em que esteja mais propício a determinada influência. A partir desta conceituação, é possível notar que o indivíduo não é vulnerabilizado pela existência de uma informação ou por sua coleta ter sido realizada, mas por serem admitidas práticas de mercado *voltadas a vulnerabilizar* e *explorar vulnerabilidades* de consumidores incapazes de *articular* o plexo de influências a que foram sujeitos quando instados a tomar determinada decisão, i.e., uma decisão de consumo ou relacionada ao exercício da cidadania.

A sociedade não ruma para o desuso de dados pessoais: o nosso estágio civilizatório é marcado pelo uso intenso da informação para diversas finalidades, e é desarrazoado supor, neste momento, uma reversão deste processo, que não é essencialmente bom, nem ruim, mas parte

da cultura (NELLIS, 1991; 2014). Por sua vez, a digitalização torna intenso o processo de virtualização e é marcada pela máxima "*humans forget machines remember*" (VILLARONGA et al., 2018). Proteger indivíduos das potencialidades de vulnerabilização e exploração requer um esforço maior de compreensão dos fenômenos inerentes aos fluxos de informação em ambientes digitais, tanto por parte dos responsáveis por políticas públicas, quanto da sociedade amplamente considerada ("titulares" de dados), não sendo mais cabível uma abordagem desleixada (JACOBS, 1961). Em todo contexto em que a vulnerabilidade puder ser explorada (CALO, 2017), o direito deverá oferecer suportes que reequilibrem de fato a relação através de mecanismos eficazes.

2.5 O padrão normativo do GDPR

Pouco após a entrada em vigor do GDPR, Zarsky (2016) avaliou os elementos fundamentais sedimentados na disciplina de proteção de dados europeia em relação às atuais práticas relacionadas a análises de dados. Observou que o GDPR entrou em vigor em um momento crucial para a economia e o ecossistema digitais, em que "substanciais riscos a direitos e liberdades estão emergindo, enquanto ao mesmo tempo vastas oportunidades para criação de valor, promoção do bem-estar e incremento de objetivos sociais diversos se desvelam" (ZARSKY, 2016, p. 996). Com isso, o erguimento de uma disciplina jurídica complexa, relativamente a fenômenos em rápida transformação, apresenta-se como desafiadora – e com muito potencial construtivo e destrutivo. Zarksy identifica a emergência do *big data* como o maior desafio da legislação, definindo o termo genericamente como se referindo a "práticas de criar e analisar vastos conjuntos de dados, que por vezes incluem informações pessoais" (ZARSKY, 2016, p. 996). Argumenta que o GDPR é "incompatível com o ecossistema de dados que a disponibilidade de big data gera", o que deverá tornar as previsões da norma rapidamente irrelevantes, ou, por outro lado, poderia levar a substanciais alterações na forma com que tais práticas são conduzidas, tornando-as subótimas e ineficientes, ao passo que limita a inovação e as utilidades disponíveis sem necessariamente prover aos cidadãos maior proteção ao seu âmbito de privacidade (ZARSKY, 2016, p. 996).

Zarsky (2016, p. 997) observa que esta incompatibilidade "provavelmente resulta de uma decisão consciente de política pública, e não de incompreensão dos legisladores europeus quanto ao que se apresenta ou mera negligência regulatória", e que o GDPR, "em geral, é

calcado em convicções filosóficas profundas relativamente à extensão com que direitos específicos, tanto de indivíduos quanto de grupos, devem ser protegidos na era da digitalização" (ZARSKY, 2016, p. 997). Busca, com esta discussão, lançar luz a possíveis discussões sobre a necessidade de reconsideração do balanceamento de interesses representado pelo GDPR e das convicções ideológicas que o sustentam, o que decorre da necessidade de se refletir acerca do real impacto do GDPR sobre práticas contemporaneamente valorosas para a sociedade, e sobre o próprio alcance do objetivo de tutelar direitos e liberdades fundamentais. Assim, o objetivo não é questionar os valores que o legislador objetivou promover, mas avaliar a consistência da forma com que identificou e endereçou os problemas a eles relacionados, e a adequação das medidas erigidas para endereçá-los.

Recentemente, foi observado que o Regulamento Europeu Geral de Proteção de Dados Pessoais (GDPR) tendeu a criar um padrão internacional normativo acerca do tema, endereçando, entre outras preocupações, àquela relacionada à fragmentariedade da *internet* quando considerada sob o aspecto territorial, e as inerentes dificuldades à aplicação dos sistemas jurídicos nacionais decorrentes desta característica (VATANPARAST, 2020). Sobretudo, argumentou-se que o GDPR foi responsável por substanciais modificações no tocante às noções acerca do controle dos fluxos de informação, dentre as quais a institucionalização da "estabilização do discurso público sobre coleta e tratamento de dados pessoais em torno da questão da privacidade (...), bem como a adoção de medidas e linguajar similares em outras jurisdições e instituições" (VATANPARAST, 2020, p. 2). Esta situação se equipara ao que Medeiros e Bygrave (2015) identificaram como efeito "*draft once, deploy everywhere"*, comum em legislações sobre tecnologia, e que denota a sua replicação irrefletida a locais e culturas diversos daquela em que foi idealizada. Ainda, que "esta estabilização discursiva tem efeitos na compreensão e enquadramento dos problemas relacionados a tratamento e dados pessoais, refletindo também uma estabilização epistemológica" (VATANPARAST, 2020, pp. 2-3, 13).

O GDPR e o padrão normativo dele decorrente foram responsáveis por realizar uma modificação no plexo normativo (VATANPARAST, 2020, pp. 13-4) relativo às atividades de tratamento de dados pessoais e aos sujeitos em torno dessas atividades. Foi argumentado que, embora atividades de tratamento de dados pessoais não sejam novas, tampouco alguma regulamentação a seu respeito, estas atividades foram anteriormente tratadas precipuamente sob a ótica da produção de conhecimento, poder, governança e infraestrutura, e hoje o são sob a ótica do direito à "privacidade", nos termos da estabilização discursiva e epistemológica

referida. Assim, pode-se dizer que as legislações voltadas a regrar atividades relacionadas ao tratamento de dados pessoais, especialmente aquelas moldadas a partir da referência europeia, tomam para si a tarefa fundamental de buscar prover equilíbrio nas relações que envolvem dados pessoais, especialmente a partir da compreensão de que ao indivíduo deve ser atribuído o poder de controlar o uso e o fluxo de dados que se lhe referem, e de tomar decisões livre de interferências reputadas indevidas, com a finalidade preponderante de resguardar espaço de desenvolvimento individual livre de potencialidades de vulnerabilização.

Porém, deve-se atentar para que o problema da coleta e tratamento de dados também pode ser avaliado sob a ótica distributiva (VATANPARAST, 2020, p. 14), questionando-se quais são os destinatários das externalidades positivas e negativas criadas pelas atividades de tratamento de dados pessoais, e ainda a partir da ótica do controle sobre conjuntos de dados, seus meios de coleta e utilização (LOI et al., 2020). O plexo normativo resultante do enfoque no direito fundamental à privacidade é responsável por erigir o titular de dados, ou seja, àquele indivíduo que pode eventualmente ser identificado por uma determinada informação, à posição jurídica em que se torna juridicamente capaz de *exercer controle* sobre tais informações, materializada em direitos como ao acesso a dados pessoais em poder de terceiros, solicitação de sua correção, de sua exclusão etc., além da atribuição de um importante e provavelmente sobrevalorizado papel, sob a ótica da efetividade, à figura do consentimento. Schwartz e Solove (2011) observam que foi no direito alemão, na década de 1970, em lei federal sobre proteção de dados, que inicialmente se igualou a informação que identifica uma pessoa daquela que *pode* identificar uma pessoa, o que continua sendo preponderante para a definição dos contextos em que seria possível este tipo de controle individual (PURTOVA, 2018).

De fato, críticas a este respeito foram feitas por pesquisadores europeus ainda na fase de formulação do GDPR, que atualizou a *Data Protection Directive* (DPD) de 1995 (EUROPEAN DATA PROTECTION SUPERVISOR, 1995). Fora argumentado que a reforma à época proposta e de fato materializada seria fundamentalmente falha por "focar estreitamente em solucionar muitos desafios relacionados à proteção jurídica com um único framework de lei de proteção de dados, divergindo da realidade das práticas de tratamento de dados pessoais do século XXI" (KOOPS, 2014). Argumentou-se que o "erro de percurso" teria sido causado pelos três *novos* objetivos da regulação de proteção de dados, a saber: (i) aumentar a efetividade do direito fundamental à proteção de dados pessoais, colocando indivíduos no controle de seus dados, particularmente no contexto de desenvolvimentos tecnológicos e globalização; (ii) aprimorar a dimensão interna [da EU] do mercado de proteção de dados por meio da redução

de sua fragmentariedade; (iii) estabelecer um *framework* amplo de proteção de dados, abrangendo todas as áreas econômicas e sociais (KOOPS, 2014). Cada um desses objetivos representaria uma falácia: a primeira relativa à ilusão de que a legislação de proteção de dados pode dar às pessoas controle sobre seus dados pessoais; a segunda falácia seria a equivocada compreensão de que a atualização normativa simplificaria a regulação sobre o tema, quando de fato tornaria o *compliance* pelos agentes de tratamento de dados, ainda que de boa-fé, complexa, improvável, custosa e pouco efetiva; e a terceira falácia seria a noção de que erigir legislação sobre proteção de dados pessoais voltada a regular *todas* as áreas da vida social em que tais atividades ocorrem seria viável ou efetivo, quando em verdade este expediente "alongaria a proteção de dados pessoais ao ponto de quebrá-la, transformando a legislação em letra morta" (KOOPS, 2014).

A respeito da abrangência da disciplina jurídica de europeia de proteção de dados pessoais, Purtova (2018) nota que aspectos do próprio GDPR e de decisões judiciais a seu respeito permitem "plausivelmente argumentar que em um futuro próximo tudo será ou conterá dados pessoais, levando a uma aplicação da [disciplina jurídica de] proteção de dados a tudo". No caso do GDPR, Purtova (2018) argumenta que se corre o risco de que os importantes valores que se visa com o Regulamento tutelar de maneira intensa vejam-se, ao contrário, em apuros, diante da sobrecarga sistêmica imposta pelo Regulamento quando confrontado com a fenomenologia inerente ao seu objeto de regulação. Por um lado, o conceito de dado pessoal é intencionalmente amplo em seu escopo material, de forma a ser capaz de acomodar ocorrências inseridas em um contexto tecnológico e social em constante mudança; por outro, este escopo material deverá crescer exponencialmente com o tempo, o que decorre de que sua aplicabilidade deverá ser atraída para cada vez mais fenômenos da vida social. Purtova (2018, p. 41) nota que este processo decore de possibilidades intrínsecas ao conceito que viabilizam a sua interpretação evolutiva; de um crescimento considerável de práticas de geração e agregação de dados; bem como avanços em técnicas de análises de dados.

A vida cotidiana e o ambiente em que a experimentamos envolve crescente mediação de nossas experiências por tecnologias da informação, o que, como consequência, leva a que tudo no meio ambiente seja ultimamente "datificado": o clima, o lixo, o esgoto, condições de saúde, preferências e hábitos diversos etc. Ao fim e ao cabo, poder-se-á razoavelmente argumentar pelo caráter de *pessoal* da maior parte, senão de todos, estes dados. Por outro lado, a não-escalabilidade do regime jurídico do GDPR e da LGPD torna este problema pernicioso (SCHWARTZ e SOLOVE, 2011). Conforme enumera Purtova (2018, p. 41), quatro

transformações principais estão no epicentro da crescente *datificação* da vida: (1) um obscurecimento da distinção entre realidade e virtualidade; (2) um obscurecimento da distinção entre humano, máquina e natureza; (3) uma reversão da escassez de informação para a abundância de informação; e (4) a virada da primazia de coisas, propriedades e relações binárias isoladas, para a primazia das interações, processo e redes (FLORIDI, 2015). Como resultado, normas com os princípios fundantes do GDPR, embora voltadas a fornecer a maior proteção jurídica sob quaisquer circunstâncias, podem na prática se mostrar de impossível cumprimento por boa parte dos destinatários e, como consequência, ignorada ou desacreditada como desarrazoada e permissiva de abusos de direitos (PURTOVA 2018, p. 41). Para Purtova (2018, p. 76), o problema identificado a partir de sua análise não está apenas no amplo escopo da disciplina jurídica de proteção de dados pessoais devido ao amplo e aberto conceito de dado pessoal; o problema jaz também na alta intensidade do *compliance* requerido pelo GDPR, indistintamente para quase todo tipo de agente de tratamento de dados (GAL e AVIV, 2020). Schwartz e Solove (2011), no contexto jurídico norte-americano, trataram extensamente desta questão, também notando a incongruência de um regime jurídico único e intenso para quaisquer atividades de tratamento de dados pessoais, o que é exacerbado pela abrangência material expansiva da norma (HALLINAN e GELLERT, 2020).

Com relação ao ideal da autodeterminação informativa, Ooijen e Vrabec (2019) observaram que, em comparação às regras anteriores sobre a disciplina vigentes na UE, o GDPR endereçou o problema em torno do controle dos indivíduos sobre dados pessoais de formas mais explícitas. Especificamente, identificaram a substancial adição à disciplina, a partir do GDPR, de diversos novos princípios com a finalidade de prover a indivíduos este tipo de controle. As pesquisadoras avaliaram a medida com que as disposições do GDPR são de fato capazes de aumentar o controle individual sobre dados que se lhe referem. Os resultados das análises permitiram concluir que os instrumentos jurídicos inseridos no regramento jurídico a partir do GDPR caminharam no sentido de prover maior controle individual sobre dados pessoais, "mas algumas ameaças ao controle individual remanescem entranhados no GDPR". Por outro lado, é possível observar que alguns direitos conferidos pela legislação têm a intenção de criar mecanismos capazes de endereçar aspectos como a invisibilidade e intangibilidade dos dados pessoais, e com isso prover a titulares efetivo controle, porém não são acompanhados de medidas ou de normas técnicas que viabilizem de fato este maior controle sobre o fluxo e o uso de informações, como é o caso do direito à portabilidade.

Sob a ótica do direito norte-americano, onde atualmente se discute a possibilidade de implementação de um padrão normativo nacional e geral sobre proteção de dados pessoais, severas críticas têm sido feitas a um possível afastamento do controle do poder e das externalidades decorrentes das atividades de tratamento de dados pessoais mediante atribuição de papel de centralidade à ideia de empoderamento decorrente da noção de que "pessoas naturais devem ter direito ao controle sobre os próprios dados" (GDPR *apud* HARTZOG e RICHARDS, 2020, p. 1734). A crítica não se relaciona ao devido poder que deve ser dado aos indivíduos em questões de seu interesse – caso em que se enquadram as situações decorrentes de tratamento de seus dados pessoais –, mas precisamente ao fato de que a atribuição de centralidade a este aspecto, frequentemente com exclusão de outras formas de controle do poder, na prática parece enfraquecer a posição dos titulares de dados e dos membros da sociedade em geral que não se beneficiam das externalidades positivas das atividades de tratamento de dados pessoais. Argumenta-se que a ilusoriedade do controle decorre do fato de que "Engenheiros projetam suas tecnologias para produzir resultados específicos. Escolhas humanas são possíveis nos limites designados pelo projeto das ferramentas que utilizam" (HARTZOG e RICHARDS, 2020, p. 1734).

Analisando extensamente a situação no Brasil, já sob a vigência da LGPD, que se argumenta, com parte da doutrina especializada, insere-se no contexto delineado de irradiação de padrões estabelecidos pelo regulador europeu sobre proteção de dados pessoais, Bioni (2018) avalia o risco de que a figura do *consentimento* assuma papel diverso daquele que têm aptidão para exercer, buscando avaliar as suas limitações e funções no plexo normativo brasileiro. O autor nota, a partir dos resultados de estudos empíricos, que é falaciosa a noção de que haveria um necessário *trade-off* em muitas das situações que envolvem disponibilização de dados pessoais em troca de determinados produtos ou serviços, e que os sujeitos estariam em geral *resignados* com a impotência em fazer prevalecer desejos relacionados à proteção de seus dados, o que decorre da assimetria de poder embutida nestas relações, reconhecendo que o problema é estrutural (BIONI, 2018, p. 218), em convergência com pesquisadores que classificaram os problemas relacionados à privacidade na contemporaneidade como uma questão arquitetônica (SOLOVE, 2006) em que a vulnerabilidade assume papel preponderante em um ciclo vicioso em que o direito à privacidade pode assumir papeis antagônicos (CALO, 2017).

No entanto, este trabalho não se relaciona a "limites do consentimento" ou da autodeterminação, importantes discussões relativas ao papel do indivíduo em consentir com o

uso dos seus dados, que deve ser entendido como elemento de contextos específicos e valorizado sobretudo a partir do conhecimento, por parte deste, de critérios decisórios utilizados com relação a si, contextualmente. Se está a falar da impropriedade da pretensão de abrangência da legislação, que contrasta com a restrição que opera com relação ao contexto (aos pressupostos) de que parte, tornando-se, desta forma, não apenas ineficiente, mas perniciosa, por favorecer a dissimulação das externalidades das atividades de tratamento de dados, a manutenção de desequilíbrios socioeconômicos e a concentração de poder, ao consolidar assimetrias de poder que se instalaram *antes da norma* pela supressão das condições de possibilidade que levaram ao surgimento de valorosas práticas envolvendo dados, ultimamente transmutando posições fáticas desiguais em posições jurídicas desiguais, com pouca possibilidade de mobilidade.

Tem sido demonstrado que legislações baseadas nos paradigmas delineados do GDPR têm representado uma dupla ineficácia: por um lado, não são capazes de tutelar o indivíduo e o livre desenvolvimento de sua personalidade, inclusive sob a ótica da saudabilidade das instituições democráticas e do controle sobre o uso relevante dos próprios dados pessoais; por outro, têm sido responsáveis por criar ou exacerbar falhas de mercado que propiciam a intensa concentração de poder observada neste setor da economia, o que paradoxalmente leva a um enfraquecimento dos direitos nacionais, e ainda a criação de barreiras de entrada artificiais em mercados que, livres de interferência, são capazes de acomodar empreendimentos com alto valor social e baixo custo, ou seja, com o potencial de propiciar desconcentração de renda e poder e oferecimento de produtos e serviços valorosos para a sociedade (inclusive voltados à tutela da privacidade).

A construção do direito à proteção de dados pessoais foi erigida a partir de uma noção de privacidade que parte de seu papel enquanto elemento central do desenvolvimento de uma esfera de autonomia individual, que adquire importância também para a construção social coletiva e política. Neste sentido, a noção de privacidade e as práticas a ela relacionadas podem ter o seu surgimento ligado a momentos históricos específicos, e pode ainda ser visto em aspectos como a arquitetura doméstica, que apenas a partir da consolidação da modernidade passou a abranger a existência de cômodos que denotavam individualidade e solitude (RYBCZYNSKI, 1986). Esta solitude é o que ultimamente permitiu ao indivíduo o desenvolvimento de ideias, conceitos e valores distintos daqueles experimentados na coletividade, e em seguida ao seu compartilhamento com a família e o grupo social, e ainda constituindo elemento essencial da cidadania. Uma vigilância voltada à conformidade social,

como aquela representada pela arquitetura ideal do panóptico (BENTHAM, 1791) tão recorrente nas ciências sociais, exerce poder precisamente a partir da supressão deste senso de solitude, pela impressão no sujeito dessa sensação permanente da possibilidade de vigilância, associada a um poder de punir a conduta tida por desviante, que idealmente leva o sujeito a se comportar conforme esperado pelo receio desta punição. Por outro lado, o fundamento do direito à proteção de dados pessoais calcado no *right to privacy* se deu em um contexto em que o conceito de informação não apresentava maiores complexidades em sua identificação e caracterização (SCHWARTZ e SOLOVE, 2011).

O enfoque na privacidade transforma os indivíduos participantes desta cultura que tem como aspecto central a intensa virtualização e digitalização em sujeitos dotados de um poder *meramente* jurídico de reclamar, com relação a dados que se reputa se lhe referem, o exercício de poder que é fração das externalidades decorrentes de tais atividades. Em outras palavras: a legislação de proteção de dados pessoais alija os *titulares* da maior parte das externalidades relativas às atividades de tratamento de dados, permitindo-lhes tão somente o direito de exercer *certo controle* sobre *certos dados* que se atribui a ele relativos, eventualmente acompanhado do desfrute de determinado serviço. É importante ressaltar que estas normas não foram e não serão capazes de alterar esta realidade fática, pois descoladas da materialidade (HALLINAN e GELLERT, 2020), e as posições jurídicas que criam relegam aos titulares de dados um papel muito pequeno no mundo atual, mantendo-os na periferia de atividades econômicas e sociais que se desenvolvem a partir de conjuntos significativos de dados, gerados pelas experiências cotidianas de inúmeras pessoas, em razão precisamente da perpetuação desta posição. Por outro lado, a própria aptidão destes mecanismos jurídicos para tutelar aspectos relacionados à subjetividade é duvidosa diante da superioridade normativa da arquitetura da rede, especialmente aliada ao poder econômico (CALO, 2013), o que tem sido demonstrado no continente europeu a partir de estudos empíricos pós-GDPR (JOHNSON et. al, 2020; JONSON et. al, 2020b; GAL e AVIV, 2020), conforme veremos adiante.

Como bem observa Zuboff (2019), o modelo de negócios que ganhou escala e significado com o Google se caracteriza pela percepção do valor que poderia ser gerado pela coleta e processamento dos dados coletados pela interação dos usuários durante a utilização do serviço de busca oferecido pela empresa, e a venda destes dados para terceiros, não envolvidos na relação original. Ou seja: a utilização dos serviços que geram dados é um aspecto *lateral* na lógica assimétrica que se constrói a partir da coleta, processamento e utilização de dados de maneira dissimulada e sem o conhecimento dos usuários. Assim, manter o foco das legislações

que visam a proteger os indivíduos e a sociedade da concentração de poder nesses serviços voltados à coleta de dados significa *comprar a distração* utilizada na origem para propiciar a acumulação original de dados e de tecnologia para processá-los e promover a manutenção do déficit epistemológico (ZUBOFF, 2020) em que se consubstancia a vulnerabilidade de titulares de dados, afastando-se de abordagens que poderiam levar à *desconcentração de poder* nesta seara.

A situação de desequilíbrio é fática, e não normativa. Argumentamos que o desequilíbrio, inicialmente, tem dois matizes: a um, se relaciona com a detenção do poder sobre a infraestrutura necessária para que as atividades informáticas ocorram; a dois, se relaciona com a incapacidade de que os indivíduos amplamente considerados se comuniquem diretamente com artefatos tecnológicos, e, com isso, compreendam-nos, o seu funcionamento e suas funcionalidades, e, ainda, de que tomem parte nos processos e práticas de fato significativas que decorrem do emprego dos frutos destes artefatos e redes. Por outro lado, a ubiquidade de atividades que envolvem tratamento de dados, e a pretensão de abrangência da disciplina, que é observada, por exemplo, na própria definição de dado pessoal, coloca este simples "mal-entendido" em posição de causar diversos danos ao sistema jurídico e à sociedade, ao passo que reforça a concentração de mercado e de poder nos mercados de dados pessoais e, consequentemente, as possibilidades de vulnerabilização e exploração de vulnerabilidades – precisamente o que se deveria combater em primeiro lugar.

Formula-se a hipótese de que é preciso reavaliar a intervenção do direito na estabilização das relações sociais que envolvem a utilização de dados pessoais de forma a endereçar estes problemas e, com isso, contar com *framework* regulatório que atinja aos fins sociais do direito. Para tanto, propomos abordagem que parta não do cânone da privacidade, que em última análise – a despeito da definição que se adote para o termo – representa apenas uma fração do que se propõe a regulamentar. A superação do paradigma da privacidade na generalidade da legislação sobre fluxos de informação deve ocorrer não apenas sob a ótica de sua *evolução* ou da superação da lógica subjetivista e do paradigma do segredo (SOLOVE, 2006), ou do estabelecimento autônomo de um direito à proteção de dados pessoais, conforme, a bem da verdade, já se encontra positivado no artigo 5º da Constituição Federal (BRASIL, 1988). O ponto que se coloca no centro da discussão é o fato de que a construção teórico-normativa da disciplina, observada em aspectos como as suas categorias fundamentais e no sistema de direitos que estrutura, é calcado nos conceitos atinentes a um direito à privacidade, o que cria incoerências internas e dificuldades interpretativas possivelmente insuperáveis,

diante do aspecto fundante destes conceitos. Os fluxos informacionais, inclusive de dados pessoais, apenas tangencial e contextualmente se relacionam com a privacidade ou a autodeterminação, pelo que estes aspectos devem fazer parte da construção da disciplina, mas não orientar a sua construção ou interpretação.

3 CONTROLE INDIVIDUAL SOBRE A INFORMAÇÃO PESSOAL

Neste capítulo, abordaremos aspectos relacionados às possibilidades de controle da informação e dos fluxos de informação por parte das pessoas a quem os dados se referem.

3.1 Dificuldades em se propiciar melhores decisões

David Lyon observa que "um aspecto-chave da emergente cultura da vigilância atual é o imperativo do compartilhar" (LYON, 2017, p. 162), que poderia ser teorizado criticamente, conforme Deborah Lupton (2015), como um meio central pelo qual as corporações monetizam o compartilhamento e a circulação do conteúdo, e criando discriminação e desvantagem para certas populações a partir do mascaramento de processos sociais e suas consequências. Observa, com Kirstie Ball (2009), que "instituições associadas a tecnologia, mídia, emprego e consumo criam uma demanda ou mobilizam recursos para focar estados psicológicos ou comportamentos íntimos" (LYON, 2017, pp. 163-4). Para Ball (2009), é preocupante a desvalorização da subjetividade na literatura sobre vigilância, que é vista comumente em termos de opressão, coerção, ambivalência ou ignorância. Ball (2009) sugere que o fato de que as pessoas não resistam ativamente e nem sequer questionem a vigilância não significa necessariamente que elas não se preocupem com isso, havendo diversas razões para que a vigilância possa ser tolerada ou mesmo desejada. Ball (2009) se baseia em estudos de McGrath (2004) sobre performatividade para explorar possiblidades de explanação de dimensões psicanalíticas da vigilância, buscando compreender, por exemplo, como a exposição pessoal é legitimada em diferentes contextos. Argumenta existir mais a se observar sobre os sujeitos que disponibilizam seus dados do que a posição reducionista e passiva em que eles muitas vezes são colocados pela literatura acerca do tema. Observa que se deve insistir no fato de que os sujeitos da vigilância

ainda fazem escolhas, mesmo que fugazes, quando solicitados pelo sistema com o qual interagem.

Os entendimentos equivocados decorrentes da centralidade do direito à privacidade na regulação da proteção de dados também alcançam o desenho de estudos empíricos sobre o tema e a interpretação de seus resultados, que, ao buscar compreender as relações entre escolhas individuais sobre disponibilização de dados pessoais e o valor atribuído por tais indivíduos à sua privacidade, frequentemente concluem no sentido de haver um "paradoxo da privacidade", "fenômeno em que pessoas dizem valorizar altamente a privacidade, porém em seu comportamento abrem mão de seus dados por muito pouco em troca, ou deixam de utilizar mecanismos para proteção de sua privacidade" (SOLOVE, 2020). A aceitação deste "paradoxo da privacidade" levou pesquisadores a buscar explicá-lo ou resolvê-lo a partir de duas linhas argumentativas: uma, que foi chamada de "argumento de valorização do comportamento", afirma que o comportamento seria a melhor métrica para avaliar como as pessoas de fato avaliam a privacidade, e, assim, os estudos demonstrariam que as pessoas, a despeito de afirmarem valorizarem a privacidade, em verdade atribuem a ela baixo valor. Por outro lado, o "argumento da distorção comportamental" entende que o comportamento não é uma métrica adequada para mensurar preferências pois seria distorcido por vieses e heurísticas, manipulação, entre outros" (SOLOVE, 2020).

Daniel Solove (2020) argumenta que o "paradoxo da privacidade" é um mito decorrente de lógica falha, pois os estudos voltados a avaliar o comportamento individual relacionado à privacidade ocorrem geralmente em contextos muito específicos, enquanto as atitudes e preocupações das pessoas a respeito de sua privacidade ou do quanto a valorizam são muito mais genéricas em sua natureza, representando um "salto lógico generalizar a partir de decisões individuais acerca de risco envolvendo dados pessoais específicos, em contextos específicos, para alcançar conclusões amplas acerca de como as pessoas valorizam a própria privacidade" (SOLOVE, 2020). Assim, o comportamento observado nos estudos sobre o paradoxo da privacidade não se relaciona com o valor da privacidade, mas, sim, decisões sobre riscos em contextos específicos. Esses contextos frequentemente envolvem dados pessoais específicos, confiados a um terceiro específico, a partir de expectativas específicas. Com isso, Solove (2020) argumenta em conclusão que decisões sobre risco são diferentes de decisões de valor. Se, por um lado, decisões sobre risco envolvem considerações sobre potenciais perdas ou danos, decisões de valor decorrem da importância subjetiva que uma pessoa atribui a algo. A este respeito, Martin e Nissenbaum (2016) observam que não se pode dizer que a privacidade de

alguém foi perdida, trocada, dada, ou violada simplesmente pela eventual cessão do controle da informação ou da própria informação, mas por isso ocorrer de forma indevida ou de forma que desborde das expectativas dos sujeitos envolvidos.

Corroborando com este argumento, vale mencionar os resultados de estudo empírico de pesquisadores da Universidade de Tilburg (KODAPANAKKAL et. al, 2020), que buscaram endereçar precisamente a limitação comum de estudos empíricos anteriores de observar avaliações ou escolhas individuais em isolamento, sem a consideração dos custos e benefícios envolvidos na decisão. Partindo do pressuposto de que tecnologias de big data apresentam custos e benefícios que influenciam na sua adoção e aceitabilidade moral, os autores observaram uma tendência de os sujeitos adotarem tecnologias com maior frequência quando seus dados são protegidos, e quando os resultados da conduta são favoráveis ao indivíduo que toma a decisão. Os pesquisadores concluíram ainda que o indutor de aceitabilidade moral mais relevante foi a proteção de dados pessoais. Assim, o estudo indica que, de fato, o *paradoxo da privacidade* é um mito, pois os sujeitos atribuem valor à proteção de seus dados pessoais ao realizar escolhas, bem como aponta para a pertinência de se avaliar a tomada de decisão a este respeito sob a ótica do risco-retorno. Ainda nesta seara, pesquisadores propuseram discussão acerca do desenvolvimento de um *direito a saber o valor monetário (preço) dos próprios dados pessoais,* como um mecanismo potencialmente capaz de ampliar o reconhecimento do indivíduo acerca do seu valor e, assim, colocá-los em melhor posição para proteger os seus dados ou para realizar trocas mais racionais a eles relacionadas, ao invés de passivamente sucumbir à apropriação de suas identidades digitais (MALGIERI e CUSTERS, 2017).

Dessa forma, uma melhor compreensão dos comportamentos individuais observados no suposto paradoxo da privacidade é em torno da ideia de risco-retorno percebido pelo indivíduo no momento de decidir sobre dados pessoais. A disciplina de proteção de dados pessoais, por sua vez, deve ser aplicável na medida em que matizes de vulnerabilidade (LUNA, 2009) estejam presentes na relação em grau considerável, afrontando direitos fundamentais. Uma forma de compreender e estruturar este critério de aplicabilidade (HALLINAN e GELLERT, 2020) é a partir da compreensão dos tipos de informação que, em dados contextos, geram potenciais de criação e exploração de vulnerabilidades (CALO, 2017), desta forma adequadamente delimitando o conceito de informação que serve de critério de aplicabilidade (HALLINAN e GELLERT, 2020). O suposto paradoxo da privacidade encontra uma de suas explicações na falta de percepção, por indivíduos que participam da chamada *cultura de vigilância*, com relação aos riscos e benefícios decorrentes de atividades de tratamento de dados, tanto em geral,

quanto nos contextos particulares em que são instados a agir. Lyon (2017, p. 157) observa que "as pessoas buscam uma forma de 'autoconhecimento' para que possam levar 'vidas melhores', ainda que apenas um pequeno fragmento dos dados seja visto por elas, e a vasta maioria termine na base de dados das corporações dos aparelhos portáteis".

O modelo europeu de disciplina de proteção de dados tem aparentemente contribuído para manter esse déficit perceptivo, pois contribui para uma relação *ingênua* das pessoas com serviços, aparelhos e ecossistemas que coletam dados, e, com isso, com o próprio mundo em que vivemos. A crença de que é possível utilizar indiscriminadamente infraestruturas, aparelhos e interfaces desenvolvidas, construídas e mantidas por terceiros sem se preocupar com a negociação inerente aos processos de distribuição, fiando-se na ficção da *autodeterminação informativa* é perniciosa e reforça o déficit epistemológico identificado por Zuboff (2020) e a própria concentração de poder, ao passo que fecha o espaço para que outras pessoas se insiram nessa economia de forma significativa, ou seja, tomem parte nas externalidades positivas das atividades de tratamento de dados. Esta constatação não pretende afirmar que indivíduos devam se resignar com qualquer tipo de exploração; ao contrário: a superação desta situação requer a capacidade de compreensão das situações em que a intervenção para a tutela de bens jurídicos seja relevante. Neste sentido, pesquisadores têm sugerido desde a edição do GDPR que se deve focar em, antes de prover uma miríade de direitos individuais a titulares de dados, construir melhores sistemas de aprendizado por máquina *ab initio*, empoderar agências que possam agir pelos interesses de titulares de dados para escrutinizar estes sistemas (EDWARDS e VEALE, 2017, p. 23), e mesmo focar em modelos de cooperativas de dados (LOI et al, 2020; BLASIMME et al, 2018) em que os titulares detém poder sobre as arquiteturas de escolha (SUNSTEIN, 2009) em que ocorrem suas decisões granulares sobre dados pessoais.

A noção de risco-retorno *percebido* pelo sujeito no momento de *escolher* parece necessitar de maior avaliação de forma a compreender tais decisões e de que forma as normas sobre proteção de dados poderiam favorecer titulares de dados quando estes são instados a realizar escolhas sobre seus dados pessoais, bem como informar o legislador no tocante às limitações relevantes ao exercício desta escolha, o que poderia indicar a necessidade de outros contornos regulatórios que visem a alcançar o equilíbrio do poder e da distribuição de externalidades decorrentes das atividades de tratamento de dados pessoais, respondendo às preocupações expostas quanto à efetividade dos padrões regulatórios hodiernos sobre proteção de dados pessoais sem cair na armadilha de uma legislação paternalista ou que limite indevidamente atividades relevantes que dependem do tratamento de dados pessoais. É

importante, neste aspecto, recordar que um *nudge* se difere de um elemento normativo precisamente em razão de não se afigurar como impositivo, e de sua superação ser possível sem um custo considerável para o destinatário; por outro lado, esta *opção padrão* representada pelo *nudge*, como decorre de noções de discussões colocadas em capítulos anteriores, exerce um papel preponderante sobre a conduta das pessoas; a sua associação a objetivos fundamentais (BEECHER-MONAS, 2007), desta forma, é desejável e pode se inserir no contexto do que Thaler e Sunstein (2003) chamaram de *paternalismo libertário*, enquanto o oposto pode representar um extremo de vulnerabilização da maior parte das pessoas, que não exerce poder sobre estas relevantes *escolhas de design*. Este tipo de preocupação é endereçado em modelos como aquele proposto por Loi et al. (2020), conforme exposto no capítulo 1 (1.2.7).

3.2 Controle individual sobre a informação pessoal

Diante da crescente complexidade das tecnologias empregadas em atividades que envolvem dados pessoais e a multiplicidade de práticas de mercado relacionadas a este tipo de atividade, tem se tornado crescentemente desafiador para consumidores deter algum tipo de controle sobre informações que, potencial ou efetivamente, se lhe referem, i.e., "seus dados pessoais". A ciência comportamental tem alertado sobre diversos riscos a esta possibilidade de controle individual, especialmente em decorrência de fatores como sobrecarga de informação e invisibilidade dos dados (KAMLEITNER e MITCHELL, 2018).

No âmbito do direito, parte-se do pressuposto legítimo de que é desejável prover aos indivíduos alguma espécie de controle sobre os próprios dados, o que decorre do *status* ou da condição de *titular* que se reputa caracterizado pela *vulnerabilidade* (CALO, 2017). O plexo de direitos em torno do que se convencionou chamar de *autodeterminação informacional,* ou *informativa*, tem sido descrito como um reflexo de valores fundamentais como autonomia, privacidade e dignidade da pessoa humana, e é ilustrado pela decisão do Tribunal Constitucional Alemão acerca do censo populacional. Na decisão, o Tribunal consignou a importância de que qualquer pessoa possa prever com suficiente grau de certeza quais informações acerca de si são conhecidas de seu círculo social, sob pena de não poder estimar com suficiente acurácia o conhecimento de que dispõem as partes com as quais se comunica, por consequência inibindo-a em sua liberdade de planejar ou decidir livremente, sem estar sujeita a influências forçadas. Com isso, descreveu a habilidade de exercer a liberdade de

decidir acerca do uso de seus dados como “controle”, definindo as balizas da autodeterminação informacional (BUNDESVERFASSUNGSGERICHT, 1983).

A Diretiva EU 2016/679, ou GDPR, é um instrumento jurídico que regula atividades de tratamento de dados no âmbito da União Europeia. É frequentemente notado por estudiosos da área que o GDPR, com relação à Diretiva anterior sobre dados pessoais vigente na União Europeia, voltou-se ao incremento das possibilidades de controle, e do controle de fato, de indivíduos sobre as informações que se reputa se lhe referem. Este objetivo é expressamente indicado nos documentos relativos à fase de elaboração do GDPR, e no texto do próprio GDPR (REDING, 2011).

Apesar deste enfoque do GDPR em prover controle individual sobre dados pessoais, a Diretiva tem sido alvo de críticas, desde antes de sua edição definitiva, por cientistas da área comportamental, relativamente à sua aparente inaptidão para endereçar as ameaças que se apresentam hodiernamente ao controle de indivíduos, i.e., a sua habilidade de exercer a liberdade de decidir (BUNDESVERFASSUNGSGERICHT, 1983) sob a ótica da forma como recebem, processam e utilizam a informação. As principais ameaças frequentemente identificadas no contexto da *economia movida por dados* podem ser caracterizadas em dois aspectos: sobrecarga de informação e invisibilidade dos dados (KAMLEITNER e MITCHELL, 2018).

OOIJEN e VRABEC (2019) partem do pressuposto de que, na missão de incrementar o controle individual, é essencial levar em conta a psicologia do processamento de informações e tomada de decisões por humanos. De fato, se considerarmos que a criação e a exploração de vulnerabilidades (CALO, 2017) relacionadas a dados pessoais na contemporaneidade têm como um de seus principais aspectos a tentativa de prever, a partir de um ou mais bancos de dados, comportamentos potenciais de indivíduos diante de determinados estímulos (GAL e AVIV, 2020; ZUBOFF, 2019), esta parece ser a única abordagem passível de resguardar o âmbito de liberdade e autonomia que outros buscam mitigar. Em outras palavras, uma das facetas da “economia movida por dados”, refere-se à utilização de conhecimento sobre o comportamento de uma pessoa ou de grupos para prever como este sujeito processará informações com que poderá ser confrontado para tomar uma decisão; assim, torna-se possível a apresentação ao indivíduo precisamente das informações que se identificou terem razoável potencial de “fazê-lo” decidir da maneira com que se deseja. É por esta razão que a capacidade de o indivíduo *perceber* o plexo de influências atuante sobre o seu processamento de informações e tomada de

decisão é fundamental para que este aja de maneira autônoma. Alguns pesquisadores têm nominado um critério de *articulabilidade* (GROARKE e PALCZEWISKI, 2016), referindo-se ao grau com que o sujeito seria capaz de *articular*, i.e., verbalizar e sistematizar, as influências (racionais e a-racionais) a que está sujeito em um dado momento, o que lhe asseguraria maior grau de autonomia.

Conceituando controle individual como "a extensão com que um indivíduo está consciente de uma situação e tem a intenção consciente e a habilidade de iniciar, terminar ou manter uma situação", OOIJEN e VRABEC (2019, p. 93) avaliaram a efetividade do GDPR em aumentar o controle individual no contexto da economia movida por dados. O objeto de estudo são os princípios introduzidos pelo GDPR na legislação europeia sobre proteção de dados pessoais que objetivam "empoderar consumidores em seu controle e avaliar a extensão com que são capazes de incrementar o controle individual sob uma perspectiva comportamental" (OOIJEN e VRABEC, 2019, p. 93). Para guiar a análise, foram identificados e caracterizados três estágios distintos em que se divide uma atividade típica de tratamento de dados legitimada pelo consentimento, especificamente a partir do ponto de vista do *titular* dos dados pessoais: (1) o estágio de recepção da informação; (2) o estágio de aprovação e uso primário da informação; e (3) o estágio de uso secundário (reuso). No primeiro, agentes que coletam dados proveem ao titular informações relativas à atividade de tratamento, por meio de uma política de uso ou documento similar. No segundo estágio, toma centralidade o contexto em que decisões sobre dados pessoais são realizadas, bem como a forma com que as solicitações para coleta de dados são realizadas, com relação aos seus efeitos sobre o controle individual. O terceiro estágio se relaciona a usos secundários dos dados em questão, e envolve a discussão acerca de "como determinadas *affordances* de dados digitalizados, como a sua intangibilidade e invisibilidade, limitam ainda mais o controle individual" (OOIJEN e VRABEC, 2019, p. 93). Para cada etapa, foram identificados os dispositivos legais do GDPR que objetivam incrementar as possibilidades de controle individual, e as maiores dificuldades que se apresentam, em um nível individual, para o exercício destas possibilidades.

No estágio de recepção da informação pelo titular de dados, o direito à explicação previsto no GDPR apresenta o potencial de endereçar o problema inerente a características como intangibilidade, invisibilidade, replicabilidade e permanência de informações em suporte digital. Porém, deixa de solucionar o problema da complexidade da informação, por assegurar tão somente uma explicação *ex ante*, i.e., genérica sobre o funcionamento de determinado(s) algoritmo(s) e não relativa ao contexto específico de tomada de decisão, não é capaz de

explicitar para o indivíduo as consequências da atividade de tratamento de dados ou de decisões automatizadas baseadas em dados pessoais. A utilização de *ícones* poderia ser bem-sucedida em mitigar problemas relacionados à complexidade e sobrecarga de informação, mas há risco de que o seu emprego simplifique a transmissão de informações a ponto de torná-la inócua (OOIJEN e VRABEC, 2019), ou de que sejam indevidamente transmitidos aos titulares sinais de controle (INNES, 2004) que proveem falsa segurança de familiaridade ou segurança (OOIJEN e VRABEC, 2019).

No estágio de aprovação e uso primário da informação, o princípio da privacidade por *design* e por padrão, e os incrementos relacionados às características do consentimento válido enquanto base legal para tratamento de dados, representam um aumento substancial das potencialidades de controle individual sobre atividades de tratamento de dados pessoais, especialmente por requererem ações afirmativas do usuário ou consumidor para que o consentimento seja válido, e ainda pela exigência de que a arquitetura dos sistemas ou interfaces que medeiam atividades de coleta de dados, por padrão, ativem as funcionalidades que resguardem em maior grau a privacidade do usuário, evitando com esses dois aspectos a superação de barreiras perceptivas conscientes do titular e o apelo a uma tomada de decisões intuitiva por parte deste. No entanto, foi verificado que em muitos casos estas medidas são estéticas e não são acompanhadas de outras que, de fato, incrementem as potencialidades de controle individual, o que se afigura necessário para que cumpram com o papel proposto (OOIJEN e VRABEC, 2019, p. 104).

No estágio de reutilização dos dados, os direitos de acesso, portabilidade e exclusão dos dados, em conjunto com os demais "direitos de controle", oferecem instrumentos aos titulares para o efetivo exercício de controle sobre seus dados. No entanto, se referem apenas a parte relativamente pequena dos dados pessoais tipicamente objeto de atividades de tratamento. Por exemplo, no caso do direito à portabilidade, apenas se incluem dados fornecidos ao agente de tratamento pelo titular, não havendo a obrigação do agente de tratamento de fornecer dados tidos por *observacionais*, como os relativos ao comportamento do usuário na interação com interfaces de dispositivos eletrônicos. O direito à exclusão, a seu turno, apenas incide sobre dados que não sejam tratados conforme base legal adequada, ou, quando a base legal é o consentimento, quando este é retirado pelo titular (OOIJEN e VRABEC, 2019). Por outro lado, as barreiras substantivas do conceito de informação criam limites fáticos à aplicação das disposições legais (HALLINAN e GELLERT, 2020).

Em seguida, trataremos das principais ameaças e possibilidades de controle individual sobre dados pessoais em cada um destes três estágios.

3.2.1 Estágio de recepção da informação

Neste estágio, agentes que coletam dados proveem ao titular informações relativas à atividade de tratamento, por meio de uma política de uso ou documento similar.

3.2.1.1 Ameaças ao controle individual

As principais ameaças ao controle individual nesta etapa são a sobrecarga de informação e a complexidade da informação.

3.2.1.1.1 Sobrecarga de informação

Um pré-requisito para que um *titular* de dados esteja no controle é que ele esteja informado acerca da atividade de tratamento de dados destacada. “Para estar no controle, um titular de dados deve ser capaz de tomar decisões que estão em linha com suas atitudes e preferências (pré-)existentes. Para tanto, o sujeito necessita de informação” (OOIJEN e VRABEC, 2019, p. 94). O estágio de “recepção de informação” é precisamente aquele momento em que os responsáveis por uma atividade que envolve coleta de dados pessoais têm a oportunidade (ou o dever) de prover ao sujeito (titular de dados potencialmente objeto de atividade de coleta) este tipo de informação. Isto pode ocorrer, por exemplo, por meio de políticas de privacidade e de uso, ou por meio de informações transmitidas ao usuário diretamente na interface de comunicação e/ou coleta dos dados, e.g., textualmente, ao lado de caixa de texto em que o usuário insere dados a serem coletados, ou de botão em que deverá clicar para aquiescer com a coleta.

Parte da doutrina tem argumentado que a detenção de possibilidades de controle individual não é factível quando o controle supostamente deve ocorrer por meio deste tipo de

informação, especialmente aquela transmitida por meio de termos e políticas de uso e privacidade (OOIJEN e VRABEC, 2019, p. 94). A quantidade massiva de informações a que indivíduos são expostos diuturnamente por meio de diversos aparelhos, mídias e serviços representa uma ameaça à habilidade e à motivação para escrutinizar os detalhes essenciais para a tomada de decisões informadas sobre privacidade. Com isso, argumentou-se que "paradoxalmente, a quanto mais informações indivíduos têm acesso sobre o que se passa com seus dados (pessoais), menos informações eles são capazes de filtrar, processar e sopesar para tomar decisões que estejam em linha com suas próprias preferências sobre privacidade" (OOIJEN e VRABEC, 2019, p. 95).

O processo decisório informado neste contexto depende de diversos fatores: em um nível elementar, primeiro, o indivíduo deve ser capaz de estimar os benefícios e sacrifícios que se podem esperar em associação à cessão dos dados, e, subsequentemente, decidir se estes estão em linha com suas atitudes, preferências e expectativas. Nos casos em que as informações estejam inseridas em políticas de uso, por exemplo, apenas após levar em conta toda a informação nelas disponível seria possível a um indivíduo conduzir esta estimativa, para então avaliar as consequências potenciais da cessão dos seus dados, e a probabilidade de que cada uma das consequências potenciais identificadas se realizem. Por fim, seria possível decidir em que medida estas consequências (sejam positivas ou negativas) estão em linha com preferências, atitudes e expectativas individuais (OOIJEN e VRABEC, 2019, p. 95). Naturalmente, isto depende, ainda, da *confiabilidade* que o sujeito pode atribuir à sua estimativa relativamente às consequências potenciais da cessão dos dados. Aparentemente, com relação à privacidade e mecanismos relacionados à proteção de dados pessoais, as partes responsáveis por prover informação aos usuários, inclusive os reguladores, partem do pressuposto de que usuários e consumidores de serviços que coletam dados são capazes de processar extensivamente toda informação a que são expostos (OOIJEN e VRABEC, 2019, p. 95), o que contraria frontalmente as evidências disponíveis a partir de numerosos e uníssonos estudos na área da neurociência acerca dos limites representados pela *sobrecarga de informações*. No contexto de rápidos desenvolvimentos tecnológicos, estes documentos por meio dos quais se provê informações aos titulares de dados estão se tornando progressivamente longos e complexos (SHORE e STEINMAN, 2015). A título ilustrativo, vale mencionar estudo conduzido na Noruega em 2016 (PALAZZO, 2016), em que se estabeleceu, a partir de estudo empírico, que o usuário médio norueguês, que possui em seu *smartphone* 33 *apps*, leva em média 32 horas apenas para *ler* os termos e condições de uso destes serviços. A forma com que

o GDPR alterou a disciplina de proteção de dados até então vigente na União Europeia aparentemente foi fator decisivo para o incremento na complexidade das informações sobre tratamento de dados transmitidas aos respectivos titulares (KOOPS, 2014).

3.1.1.1.2 Complexidade da informação

As limitações relativas ao fenômeno da sobrecarga de informação se referem a habilidades cognitivas em geral, i.e., indistintamente com relação a quaisquer características pessoais ou grupais do receptor, ou do tipo de informação transmitida. No entanto, pesquisadores têm identificado outro tipo de limitação à plena compreensão de informações relacionadas à proteção de dados, que se relaciona com o papel de diferentes níveis de conhecimentos específicos ou literacia do receptor da informação (HARGITTAI, 2007). Park (2013) avaliou a capacidade de consumidores compreenderem aspectos relacionados à uma política de privacidade "comum" e observou que em média os participantes acertaram apenas duas de sete perguntas simples (e.g., se um *website* é legalmente autorizado a compartilhar informações sobre a pessoa com parceiros sem lhe transmitir o nome destes parceiros) sobre a proteção de seus dados. Jensen e Poots (2004) compararam o grau de legibilidade de 64 políticas de privacidade de empresas americanas (e.g., Google, eBay) para públicos com diferentes níveis de escolaridade, e determinaram que apenas 6% das políticas eram legíveis, i.e., suficientemente compreensíveis, aos usuários da *internet* com no máximo o ensino médio completo; 54% das políticas estavam *além* das possibilidades de compreensão dos usuários com mais de 14 anos de escolaridade, e 13% das políticas estavam *além* da compreensão também dos usuários com pós-graduação. Estes resultados indicam não apenas que as políticas não são aptas à finalidade a que se propõem, i.e., informar o público em geral, mas também que são *extremamente* complexas ou de outra forma equivocadas na realização de seu intento comunicativo, pois são incompreensíveis para parcela considerável de usuários altamente escolarizados. OOIJEN e VRABEC (2019, p. 96) observam um efeito de "severa ruptura no controle da informação" nos ambientes de rede contemporâneos, que trouxeram novos matizes à capacidade humana restrita de processamento de informações, ora levados a uma pressão extrema: a assimetria de conhecimento entre os responsáveis pelos sistemas que coletam e processam dados se agrava pela sofisticação dos algoritmos e inteligências artificiais projetadas para desempenhar atividades com dados, e ainda pela complexidade dos fluxos de compartilhamento e replicação destes dados, o que leva frequentemente ao desconhecimento,

por parte do próprio agente de tratamento de dados, com relação a um eventual receptor destes dados inicialmente coletados (OOIJEN e VRABEC, 2019, p. 96).

3.2.1.2 Formas com que o GDPR endereça os problemas deste estágio

O GDPR endereça os problemas identificados neste estágio a partir do direito à explicação e disposições acerca de ícones.

3.2.1.2.1 Direito à explicação

OOIJEN e VRABEC (2019, p. 96) identificaram três artigos do GDPR voltados a incrementar as possibilidades de controle individual durante a etapa de recepção da informação. Todos são relacionados ao que foi chamado de "direito à explicação", termo cunhado por Goodman e Flaxman (2016 *apud* OOIJEN e VRABEC, 2019, p. 96) em referência a decisões automatizadas, que deverão ter a sua ocorrência informada ao titular, bem como deverão ser reveladas "informações significativas sobre a lógica envolvida [no processo automatizado de tomada de decisões], bem como a sua significância e a consequência objetivada" (GDPR, Article 15). O GDPR, assim como o LGPD, também requer que o responsável por uma atividade de tratamento de dados revele ao titular informações como a própria ocorrência da atividade, os propósitos da atividade de tratamento, a identidade do controlador e o período de duração da atividade. No entanto, "direito à explicação" recebeu (e ainda recebe) grande atenção de pesquisadores e formuladores de políticas públicas pelo seu suposto potencial de levar aos titulares de dados verdadeira *compreensão* acerca *do que* é feito com os seus dados e as potenciais consequências desta utilização. Desta forma, poderia significar um avanço para com relação às possibilidades de controle individual sobre dados pessoais caso oferecesse respaldo para a avaliação destas potenciais consequências, e, com isso, parâmetros para um melhor e mais informado processo de tomada de decisão pelo titular dos dados. No entanto, o GDPR não é explícito com relação a quão específica deverá ser esta explicação, tampouco quanto aos seus limites. Estes aspectos têm sido alvo de intensos debates acadêmicos, que

buscam compreender que tipo de construção teórica em torno deste "direito à explicação" pode contribuir para o incremento na autonomia dos titulares[6].

Neste sentido, Wachter, Mittelstad e Floridi (2017) avaliaram as potencialidades deste direito, dada a sua literalidade no GDPR, as discussões de políticas públicas ao tempo de sua edição, a forma como o tema era tratado na legislação e jurisprudência europeias anteriormente, e o objeto sobre o qual a norma deve atuar para prover aos titulares proteção contra a vulnerabilização e os direitos específicos que prescreve. Este direito é por vezes visto como mecanismo promissor para prover maior transparência e *accountability* ao emprego de algoritmos, inteligência artificial, robótica e outros sistemas automatizados por governos e empresas, considerando que por vezes tais sistemas apresentam efeitos inesperados ou não-intendidos, e se estruturam de forma complexamente organizada. Neste ponto, podemos observar que os tais mecanismos não são incompreensíveis em sua operação, mas precisamente voltados a organizar a informação de maneiras que o ser humano não poderia fazer. Nos contornos delineados por Hallinan e Gellert (2020) ao analisar o papel do conceito de informação, podemos notar que alguns algoritmos, por exemplo, são voltados a obter informação, ou conhecimento, a partir de dados que não necessariamente representariam, caso se apresentasse a necessidade de cognição humana de significado para a sua caracterização, como informação – ao menos não antes do emprego da ferramenta. Vemos com isso, ainda, o tipo de incompatibilidade apontado por Zarsky (2018) e diversos outros, de pressupostos e princípios como o da minimização e finalidade especificada *ex ante* para atividades de tratamento de dados pessoais com as atividades realizadas contemporaneamente com dados – uma vez que, muitas vezes, o papel de tais algoritmos é precisamente o de extrair conhecimentos de dados que o ser humano possivelmente não seria capaz. Neste sentido, vemos o substrato fático do argumento de Wachter et. al (2017) quanto à inexistência de um direito à explicação, conforme popularmente concebido, no GDPR, que inclui ainda argumentos decorrentes de análise normativa. No mesmo sentido, Edwards e Veale (2017) argumentaram que, embora atrativo, um direito à explicação não é contemplado pela redação normativa do GDPR, e, ainda que o fosse, o tipo de explicação capaz de *endereçar a lógica* do sistema, no sentido da forma com a qual ele chega de uma informação a outra, aprioristicamente, provavelmente não é compatível com as técnicas computacionais hodiernas.

[6] Cf. EDWARDS e VEALE, 2017; WACHTER et al., 2017.

Wachter et. al (2017, p. 6) notam que, ao referir-se a um direito à explicação (de decisões automatizadas), pode-se estar a falar de dois tipos de explicação, a saber: por um lado, uma explicação referente à *funcionalidade do sistema*, i.e., quanto à lógica, ao significado, às consequências antevistas e a funcionalidade geral de um sistema de tomada de decisões automatizado. Assim, englobaria aspectos como a especificação dos requisitos de funcionamento do sistema, árvores decisórias, modelos pré-definidos e estruturas classificatórias. Por outro lado, pode-se referir a um direito à explicação de *decisões específicas*, i.e., a *rationale*, os motivos, as circunstâncias individuais e específicas de uma decisão automatizada. Inclui, por exemplo, o sopesamento de atributos realizado e o seu resultado, regras decisórias pré-definidas para determinadas circunstâncias informações sobre referência a grupos de referência ou perfis. Também pode-se distinguir um possível direito à explicação quanto ao *momento* em que ocorreria a explicação, com relação ao processo de tomada de decisão em questão. Assim, pode-se falar, por um lado, de uma explicação *ex ante*, que ocorre antes da ocorrência da tomada de decisão automatizada. Este tipo de explicação, por decorrência lógica, apenas pode abarcar o primeiro tipo de direito à explicação acima referido, qual seja, aquele referente à funcionalidade do sistema, uma vez que a decisão específica ainda não ocorreu, e com isso não se pode falar de seus motivos ou resultado do sopesamento de atributos. Ainda quanto ao momento, a outra categoria é a de uma explicação *ex post*, ou seja, aquela que ocorre após uma tomada de decisão automatizada ter ocorrido. Uma explicação *ex post* pode abarcar explicações tanto quanto à funcionalidade do sistema, quanto relativamente a decisões específicas.

Quanto às limitações do direito conforme consubstanciado no GDPR, Wachter et. al (2017, p. 27) observam que, aparentemente, algumas legislações nacionais europeias que eram baseadas na Diretiva 95/46/CE conferiam maior proteção a titulares de dados relativamente a explicações de tomadas de decisões automatizadas. Os pesquisadores atribuem este fato ao menos abrangente direito de acesso do GDPR (Article 15, comentários no Recital 63) em comparação com o previsto na Diretiva 95/46/CE, considerando que este é um dos três direitos elencados no GDPR que se associam mais diretamente ao chamado direito à explicação – os outros dois são as obrigações do controlador de notificar um titular de dados sobre a ocorrência de uma atividade de tratamento de dados pessoais (Articles 13-14 e comentários nos Recitals 60-62) e salvaguardas contra a adoção de processos automatizados de tomada de decisões (Article 22(3), comentários no Recital 71) (WACHTER et. al, 2017, p. 7). Argumentam Wachter et. al (2017, p. 28) que o GDPR, mediante a utilização de um tipo de semântica

orientada para o futuro, bem como da sobreposição terminológica com obrigações de notificação, intende limitar ainda mais o direito de acesso relativamente à tomada de decisões automatizadas, e a apenas assegurar o direito a explicação quanto à funcionalidade do sistema, ou seja, explicações genéricas, em contraposição a um direito à explicação que envolve o esclarecimento da operação específica que levou à tomada de decisão individual. As interpretações legais no continente tendem, ainda, a conformar este direito à explicação com aqueles direitos eventualmente titularizados pelo agente de tratamento de dados, especialmente aqueles relativos à proteção da propriedade intelectual, neste caso, relativamente aos sistemas de decisão automatizada que se utilizam de dados pessoais, e a cujo respeito existe algum tipo de direito à explicação (WACHTER et. al, 2017).

Compreender a informação recebida e/ou disponível é o que conduz a maior controle. A propósito dos valiosos debates acadêmicos sobre o tema, pode-se dizer que o GDPR provê apenas um direito à explicação que tem sido chamado de *ex ante* (WACHTER et. al, 2017): uma explicação genérica acerca da funcionalidade de determinado algoritmo que processa informações para chegar à determinada conclusão, i.e., decidir de forma automatizada a partir de dados. Neste sentido, Edwards e Veale (2017) argumentaram que a exploração acerca de um tipo de explicação "sujeito-cêntrica" e focada em aspectos particulares de um modelo pode ser promissora na sua capacidade de levar ao sujeito ao qual se refere uma decisão automatizada algum grau de compreensão acerca da forma com que determinados aspectos contribuem para o processo de tomada de decisão pelo sistema. Ainda assim, argumentam que a ideia de um "direito à explicação" se afigura como uma distração, representando "uma nova forma de falácia da transparência" (EDWARDS e VEALE, 2017), e que é mais útil o emprego de direitos factíveis de gerar efeitos no caso concreto mais diretamente, como o direito ao apagamento de dados pessoais e o direito à portabilidade, por um lado, e a privacidade por *design,* aliada a práticas éticas e passíveis de controle externo, por outro (EDWARDS e VEALE, 2017).

Por outro lado, um direito à explicação *ex post* significaria a possibilidade de o titular exigir informações sobre a lógica e as circunstâncias específicas aplicados a uma determinada decisão, os dados ou os aspectos que foram considerados neste caso particular, e qual peso foi-lhes atribuído na árvore decisória ou outra estrutura classificatória empregada (OOIJEN e VRABEC, 2019, p. 97). Apenas uma explicação do tipo *ex post* seria capaz de endereçar o problema da assimetria informacional entre os sujeitos envolvidos, i.e., agente de tratamento e titular de dados, pois adentraria no nível particular da decisão (OOIJEN e VRABEC, 2019, p. 97). A seu turno, no campo da linguística, tem sido demonstrado que o uso de explicações

menos concretas resulta em maiores dificuldades no processamento das informações recebidas, como reduzidas compreensão da linguagem e retenção da informação (HOLMES e LANGFORD, 1976 *apud* OOIJEN e VRABEC, 2019, p. 97), resultando em pouca efetividade de explicações *ex ante* no contexto delineado, embora dificilmente se possa dizer que não constitua uma espécie de avanço, ainda que pequeno, com relação às capacidades potenciais de um indivíduo compreender o contexto em que cede os seus dados e, com isso, exercer sobre eles maior grau de controle.

3.2.1.2.2 Ícones

O GDPR não especifica como o dever de informação, correlato ao direito de acesso assegurado aos titulares de dados, deve ser cumprido pelos agentes responsáveis por atividades de tratamento de dados pessoais (OOIJEN e VRABEC, 2019, p. 97). No contexto da *web* 2.0, a informação completa é comumente disponibilizada por meio de termos e condições de uso, ou em políticas de privacidade. No caso do uso de serviços, aplicativos e *websites* em dispositivos móveis, o problema é exacerbado diante da impraticabilidade da prática da leitura de longos documentos. Ooijen e Vrabec (2019) observam que uma potencial saída para a falência representada pela disponibilização de informações primariamente pelos referidos documentos seria a utilização de *ícones*, que são caracterizados como "imagens padronizadas que transmitem informações chave sobre a atividade de tratamento de dados" (OOIJEN e VRABEC, 2019, p. 97). O seu potencial em incrementar as potencialidades de controle individual está fundado em dois aspectos: primeiramente, ícones simplificam a compreensão da informação transmitida; em segundo lugar, elas representam grande economia de tempo para o receptor da informação, em comparação à necessidade de ler textos escritos (OOIJEN e VRABEC, 2019, p. 97). Assim, são instrumentos aptos a endereçar o problema da *sobrecarga de informação* no âmbito do processo decisório sobre cessão de dados pessoais. O artigo 12(7) do GDPR dispõe sobre a opção de se utilizar ícones padronizados para prover, de forma claramente legível, inteligível e significativa informações sobre a atividade de tratamento de dados que se pretende realizar (OOIJEN e VRABEC, 2019, p. 97).

Ooijen e Vrabec (2019, p. 98) argumentam que ícones apresentam potencial de endereçar os três problemas relacionados à disponibilização de informação sobre atividades de tratamento de dados previamente identificados: primeiro, pois reduzem drasticamente a carga

de informação transmitida; em segundo lugar, pois reduzem a complexidade da informação; em terceiro lugar, e como resultado da menor carga e menor complexidade de informações, menos tempo e atenção são requeridos para que consumidores e usuários apreendam as implicações de ceder os seus dados, e, com isso, exerçam maior grau de controle sobre eles ao serem instados a tomar decisões a seu respeito. Por outro lado, a utilização de ícone padronizados teria o potencial de reduzir problemas relacionados a influência indevida exercida sobre usuários por sinais gráficos não padronizados ou logomarcas que possam indevidamente transmitir sinais de confiabilidade (HOOFNAGLE e URBAN, 2014). De fato, foi identificado que titulares de dados têm maior propensão a consentir com o uso de seus dados por grandes e conhecidas empresas, em comparação a solicitações por empresas menores ou menos conhecidas (GAL e AVIV, 2020). Por outro lado, *sinais* são capazes de gerar diferentes respostas perceptivas em seus destinatários, frequentemente com um efeito *enquadratório* do processo de tomada de decisão que exerce papel relevante na determinação de como as informações são interpretadas, do valor e peso que lhes são atribuídos, e mesmo quanto a quais informações adentram a esfera de consciência e quais permanecem fora do âmbito de consideração consciente (INNES, 2004). A utilização de ícones padronizados voltados a claramente comunicar riscos associados à cessão de dados pessoais poderia tornar os usuários menos propensos a influências irracionais oriundas de sinais triviais ou empregados de forma deliberadamente enganosa.

Por outro lado, ícones não são capazes de transmitir informações de forma ampla e detalhada, mas apenas de forma geral e simplista (OOIJEN e VRABEC, 2019, p. 98). Desta forma, a sua utilização apenas se justifica caso seja voltada a tornar os usuários conscientes do contexto em que se insere o seu processo decisório, não devendo servir para dar aos usuários a sensação de que estão recebendo, através deles, *toda* a informação relevante à atividade de tratamento de dados. Conforme observa Nissenbaum (2011), na economia movida por dados são os detalhes intangíveis, mais do que as informações genéricas, que guardam maior significância. Este tipo de comunicação poderia ser confundido como um *sinal de controle* (INNES, 2004), dando aos receptores da informação uma falsa sensação de segurança decorrente de aparente abrangência e suficiência das informações transmitidas e da proteção legal conferida pela disciplina de proteção de dados pessoais. Estes riscos devem ser considerados na implementação de ícones na comunicação entre agentes de tratamento e titulares de dados.

3.2.2 Estágio de Aprovação e Utilização Primária

Neste estágio, toma centralidade o contexto em que decisões sobre dados pessoais são realizadas, bem como a forma com que as solicitações para coleta de dados são realizadas, com relação aos seus efeitos sobre o controle individual.

3.2.2.1 Ameaças ao controle individual

Nos casos em que o consentimento é necessário e solicitado do titular para a coleta e tratamento de dados que se lhe referem, observou-se que mudanças sutis no contexto em que o consentimento é requerido podem afetar inconscientemente o processo de tomada de decisões de indivíduos a este respeito, o que poderia levar a uma perda de controle individual. De fato, no contexto da economia movida por dados, as partes com o dever de informar sobre práticas de tratamento de dados eventualmente empregam deliberadamente técnicas que induzem consumidores e usuários a processar as informações recebidas de forma intuitiva, especialmente a partir de sutis diferenças na *arquitetura de escolha* (OOIJEN e VRABEC, 2019, p. 98). Fatores contextuais que incentivam comportamentos relacionados à cessão de dados pessoais são economicamente atrativos para os agentes que coletam dados, e desta forma algumas interfaces podem funcionar em favor destes agentes, atuando fora do âmbito de consciência do consumidor ou usuário que interage com a interface e, com isso, apresentado o potencial de colocá-lo em posição vulnerável (CALO, 2013). Um exemplo elementar deste tipo de influência da arquitetura de escolha pode ser vislumbrado no âmbito das "opções padrão": diversos estudos indicaram que usuários tem uma maior propensão a escolher as opções indicadas por padrão em múltiplos contextos, de doação de órgãos (JOHNSON e GOLDSTEIN, 2003) a cessão de endereço de e-mail (JOHNSON et al., 2002). Aparentemente, diversas razões se somam ou se alternam para explicar essa ocorrência: por um lado, indivíduos muitas vezes não leem ou não compreendem as informações recebidas no contexto de tomada de decisão (OOIJEN e VRABEC, 2019), e com isso tenderiam a manter a opção indicada por padrão; por outro, foi demonstrado que indivíduos frequentemente interpretam inconscientemente as opções padrão como recomendações, por demonstrarem um viés inconsciente para com o *status quo* (MCKENZIE et al., 2006), e porque têm dificuldade em rejeitar uma opção indicada por

padrão devido à aversão à perda (SMIITH et al., 2013). No contexto do processamento intuitivo de informações, estes efeitos se tornam mais acentuados (OOIJEN e VRABEC, 2019, p. 99).

Para além de afetar o exercício do controle em si durante a tomada de decisões, elementos presentes no contexto também afetam as experiências subjetivas acerca do controle, afetando a forma com que indivíduos percebem o controle (OOIJEN e VRABEC, 2019, p. 99). Brandimarte et al. (2013) demonstraram que um incremento na percepção de controle de um indivíduo – mas não do controle de fato – com relação à cessão e permissão de acesso a seus dados pessoais aumenta a sua tendência a revelar informações sensíveis para um público mais amplo. Hoofnagle e Urban (2014) observaram que 62% dos indivíduos que responderam a um questionário acreditavam que a mera existência de uma política de privacidade em um *website* implicava no fato de que este *website* não estaria autorizado a compartilhar as suas informações sem permissão. É provável que a presença de indícios contextuais que sinalizam controle, como a mera existência e publicização de uma política de privacidade, reduza as preocupações individuais com a privacidade e potencialize comportamentos tendentes à cessão de dados (OOIJEN e VRABEC, 2019, p. 99). De fato, estas observações e conclusão são consistentes com achados acerca do papel de *sinais* no processo perceptivo e interpretativo de situações de risco, segurança ou controle (INNES, 2004). Com isso, "a comunicação a indivíduos por meio de arquiteturas de escolha e *design* deve ser sempre acompanhada de uma situação que de fato incremente o controle" (OOIJEN e VRABEC, 2019, p. 99)., reforçando o papel da arquitetura e do arquiteto de escolhas na incorporação de determinados valores às práticas sociais (SUNSTEIN, 2009).

Em suma, aspectos contextuais e arquitetônicos se relacionam ao controle de diversas formas, das quais se destacam duas: primeiro, aspectos sutis são capazes de afetar preferências pessoais acerca de privacidade por meio da exploração de limitações à racionalidade estrita e apelo ao processamento intuitivo da informação pelos indivíduos instados a produzir e ceder dados. Em segundo, aspectos contextuais são capazes de aumentar ou reduzir a percepção individual de controle, o que, a seu turno, também exerce efeitos sobre expectativas relacionadas à privacidade e afeta o processo racional de tomada de decisões a este respeito.

3.2.2.2 Formas com que o GDPR endereça os problemas deste estágio

O GDPR buscou endereçar os problemas relacionados ao controle individual no estágio de aprovação e uso primário da informação a partir, sobretudo, de dois elementos: o consentimento (e suas condições especificadas); e o princípio geral da privacidade por padrão.

Quanto ao consentimento, vale destacar que, de forma genérica, qualquer arquitetura que exija ação ativa do usuário, para além de mero silêncio ou inação, deverá contribuir em particular medida para incrementar o controle individual, pois aumenta as possibilidades de um processo reflexivo pleno, que leve em consideração todos os aspectos envolvidos na cessão de dados pessoais em questão. Neste sentido, a vedação do GDPR ao emprego de caixas de seleção pré-marcadas com o intuito de coletar consentimento do usuário representa um incremento às possibilidades de controle individual (OOIJEN e VRABEC, 2019, p. 100). Por outro lado, o GDPR fez uma distinção entre "consentimento inequívoco (não-ambíguo)" e "consentimento explícito". O último foi exigido apenas no contexto de consentimento para o tratamento de dados sensíveis. O fato de que, para dados pessoais em geral, o GDPR exigiu "consentimento inequívoco" pode representar uma maior possibilidade de influência indevida no processo decisório, por dificultar a segregação de informações relevantes ao contexto. Por exemplo, uma interface pode mostrar para o usuário uma caixa de texto, ao lado da informação "Insira o seu endereço de e-mail para receber ofertas"; ao inserir o seu endereço de e-mail e enviá-lo por meio do formulário, o usuário consentirá legalmente à utilização do dado pessoal compartilhado, pois o consentimento foi "inequívoco". A eventual exigência de consentimento explícito representaria a impossibilidade desta ocorrência; seria necessário que o titular aquiescesse com uma frase do tipo "Eu consinto a..." (OOIJEN e VRABEC, 2019, p. 100).

Para incrementar as possibilidades de controle individual no estágio de aprovação e uso primário do dado, o GDPR também traz o princípio da privacidade por padrão em seu Artigo 25. Com isso, o agente de tratamento de dados deverá respeitar comando de proporcionalidade nas atividades de tratamento de dados pessoais, coletando e de outras formas utilizando tão somente aqueles que sejam necessários e adequados à finalidade especificada. Além disso deverá implementar medidas técnicas e organizacionais adequadas para assegurar o cumprimento deste comando, e ainda para assegurar que, por padrão, dados pessoais não sejam tornados acessíveis ao público amplamente considerado sem a intervenção do titular dos dados. Na prática, isso significa que, por padrão, em tese, as opções mais favoráveis à preservação da privacidade e dos dados pessoais de usuários e consumidores deverão ser aquelas ativadas por padrão, e qualquer modificação deverá ser ativa e deliberadamente realizada pelo usuário, i.e.,

ao se cadastrar em uma rede social, por padrão o perfil do usuário não deverá ser acessível publicamente.

3.2.3 Estágio de controle e (re)uso dos dados

Este terceiro estágio se relaciona a usos secundários dos dados em questão, e envolve a discussão acerca de "como determinadas *affordances* de dados digitalizados, como a sua intangibilidade e invisibilidade, limitam ainda mais o controle individual" (OOIJEN e VRABEC, 2019, p. 93).

3.2.3.1 Ameaças ao controle individual

No contexto da coleta e utilização iniciais da informação, o controle individual pode ser exercido de forma mais factível e objetiva, pois as razões para a atividade de tratamento de dados são mais específicas e próximas do titular, i.e., criar uma conta para utilizar um serviço online (OOIJEN e VRABEC, 2019, p. 101). Porém, a utilização de dados pessoais no contexto da economia moderna dificilmente ocorre apenas uma vez: usualmente, ele será compartilhado e reutilizado múltiplas vezes. Neste estágio de reutilização dos dados surgem, em tese, cada vez mais ameaças ao controle individual sobre dados pessoais, especialmente em decorrência da opacidade frequentemente presente nas atividades realizadas, do seu afastamento do titular de dados pessoais, e da redução que se opera pela efetiva utilização da informação extraída dos dados em substituição a eventual solicitação direta ao titular.

O crescente número de ameaças ao controle individual neste terceiro estágio guarda estreita relação com características da informação *digitalizada*: intangibilidade, invisibilidade e escopo do fluxo são características que limitam as possibilidades de exercício de controle individual sobre "um dado pessoal". Especialmente na terceira etapa estas características tornam-se problemáticas, pois a tendência da informação digitalizada é se espalhar pelo ambiente digital (OOIJEN e VRABEC, 2019, p. 101). A intangibilidade, ou seja, a inexistência em suporte físico específico, do dado pessoal, representa uma dificuldade inerente para o exercício de controle sobre este "objeto" (KAMLEITNER e MITCHELL, 2018; HALLINAN e GELLERT, 2020). Até mesmo a percepção de valor atribuída em um nível individual

aparentemente é maior para com relação a objetos físicos do que com relação a seus equivalentes digitais, i.e., uma fotografia em suporte físico ou em suporte digital (ATASOY e MOREWEDGE, 2017), ou para bens com relação aos quais se conheça ou possa atribuir valor monetário (MALGIERI e CUSTERS, 2017).

A invisibilidade, a seu turno, representa o fato de que, para consumidores e usuários, os dados que se lhe refiram ou à sua atividade e que são coletados por terceiros permanecem, para estes usuários ou consumidores, invisíveis, o que é preponderante com relação aos dados comportamentais (OOIJEN e VRABEC, 2019, p. 101) e inferências deles decorrentes. Estudos indicaram que tornar os dados de alguma maneira visíveis aos usuários os poderia auxiliar a compreender o que ocorre com seus dados pessoais, e os faria sentir em melhor posição para exercer controle sobre tais dados (OOIJEN e VRABEC, 2019; NIEZEN et al., 2010; KAMLEITNER e MITCHELL, 2018).

Por fim, a abrangência do fluxo posterior dos dados pessoais representa que, após sua coleta e tratamento iniciais, eles frequentemente são conduzidos para outros agentes de tratamento de dados ou mesmo novos controladores (OOIJEN e VRABEC, 2019, p. 101). Em muitos casos, isto envolve a restruturação ou fusão de bancos de dados que antes estavam em poder de partes distintas, ou mesmo a fusão de bancos de dados próprios ou adquiridos de terceiros com bancos de dados disponíveis publicamente ou oriundos de um conjunto de dados colaborativo entre operadores de dados (OOIJEN e VRABEC, 2019, p. 101). Os fluxos pelos quais os dados pessoais são conduzidos neste processo – vale dizer, se replicando, e não se "movendo" – são usualmente complexos e opacos, e em muitos contextos podem fugir do âmbito de conhecimento dos próprios agentes que iniciaram o tratamento de determinados dados pessoais (OOIJEN e VRABEC, 2019, pp. 101-2).

3.2.3.2Formas com que o GDPR endereça os problemas deste estágio

O GDPR concede a titulares de dados uma série de "direitos de controle" que adquirem importância por seu potencial em prover controle individual sobre os dados pessoais nos estágios posteriores à coleta e utilização primária (OOIJEN e VRABEC, 2019, p. 102). Destacaremos como pesquisadores têm caracterizado, neste contexto, os direitos de acesso aos dados objeto de tratamento e o direito à portabilidade dos dados. Outros "direitos de controle" são o direito de ter os dados pessoais excluídos de determinado banco de dados, e de retificar dados incorretos ou desatualizados.

3.2.3.2.1 Direito de acesso e portabilidade

O direito de acesso a dados objeto de atividade de tratamento visa a garantir que o titular possa receber uma cópia dos dados que se lhe referem e que estejam sob tratamento por parte de um determinado agente. Desta forma, o titular poderá tomar conhecimento destas atividades e de suas características essenciais, como quais dados são objeto da atividade, qual o período de duração da atividade e qual é o propósito da atividade, e com isso poderá verificar a sua legalidade e solicitar providências que entenda necessárias. O direito à portabilidade, a seu turno, assegura que titulares de dados possam obter uma cópia de seus dados pessoais em poder de determinado agente de tratamento, com a finalidade precípua de entregá-los a outro agente de tratamento de dados ou utilizá-los para finalidades próprias, intendidas e realizadas pelo próprio titular (OOIJEN e VRABEC, 2019, p. 102).

O mercado de serviços prestados em ambientes digitais, e.g., redes sociais, comércio eletrônico, *streaming* de mídia, é marcado por intensa concentração, que, em alguns casos, pode conduzir a incentivos para a instauração de monopólios. Um direito à portabilidade poderia reduzir os incentivos para a concentração de mercado ao ampliar as possibilidades de que dados relevantes para determinadas atividades econômicas sejam transmitidos a um ou outro controlador, a critério do próprio titular, ou ainda ao permitir que o próprio titular faça uso do conjunto de dados (EUROPEAN DATA PROTECTION SUPERVISOR, 2015). Assim, um dos principais objetivos do direito à portabilidade dos dados pessoais, e possivelmente o mais promissor, é fortalecer o controle de um titular sobre os dados que se lhe referem, e propiciar que estes titulares, de formas mais diretas, tenham acesso ao valor gerado a partir de seus dados pessoais – potencialmente superando a categoria jurídica de "titulares" para tornarem-se cocriadores daquilo que depende de seus dados pessoais e influindo diretamente em seu armazenamento e utilização.

Previsões legais relacionadas a um direito à portabilidade, em alguma medida, endereçam as características de invisibilidade e intangibilidade dos dados pessoais em suporte digital, especialmente na fase de reutilização dos dados pessoais. Conforme observaram Kamleitner e Mitchell (2018), a visualização dos fluxos de dados pode ser útil para ampliar a compreensão de indivíduos acerca do que é feito com dados que se lhe referem após o consentimento. Pesquisadores têm sugerido, ainda, a criação de plataformas por meio das quais,

de forma segura, agentes de tratamento disponibilizem dados pessoais sob sua guarda aos titulares a que se referem, que poderiam desempenhar diversas tarefas a partir deste sistema de *visibilização* dos dados (LOI et. al, 2020).

No entanto, dois importantes aspectos limitam consideravelmente estes potenciais: primeiramente, dados observados, em oposição a dados que são inseridos pelo próprio usuário, não são em regra passíveis de solicitações de portabilidade, e, assim, não comporão bancos de dados recebidos em decorrência deste tipo de solicitação do titular. Com isso, alguns dos dados mais invisíveis no âmbito da economia movida por dados – como os dados de localização e comportamentais – deixam de ser abrangidos por esta ampliação do controle do titular sobre os próprios dados. O segundo aspecto limitador se refere ao já mencionado aspecto de que o direito à portabilidade representa o direito de o titular receber uma *cópia* do banco de dados em poder de um agente de tratamento, mas não conduz à exclusão automática dos dados do banco de dados original. Para que isso ocorra, é necessário que o titular exerça outro direito, qual seja, o direito à exclusão dos dados, que deverá observar os seus próprios requisitos legais (OOIJEN e VRABEC, 2019, p. 103), e, ainda, sujeitar-se aos aspectos próprios da materialidade sobre a qual objetiva atuar (HALLINAN e GELLERT, 2020; VILLARONGA et. al, 2018).

3.2.3.2.2 Direito de exclusão dos dados

O direito de exclusão de dados pessoais é direito assegurado pelo GDPR a titulares de dados nos casos em que a atividade não seja conforme à legislação, como nas hipóteses em que o dado não mais seja necessário à persecução da finalidade para o qual fora coletado, ou quando o titular retira o consentimento, e este era a base legal para a atividade de tratamento (OOIJEN e VRABEC, 2019, p. 103). Um aspecto importante do direito à exclusão reside no fato de que um controlador que recebe tal solicitação de um titular, e que esteja obrigado a cumpri-lo, deverá notificar quaisquer agentes com os quais o dado tenha eventualmente sido compartilhado para que procedam com a exclusão de *cópias* que eventualmente possuam, exceto caso isso seja impossível ou envolva um esforço desproporcional (OOIJEN e VRABEC, 2019, p. 103). Esta obrigatoriedade guarda dificuldades práticas importantes, frequentemente insuperáveis, o que gera distorções nos mercados de dados (GAL e AVIV, 2020).

Quanto a este direito, cabe relembrar algumas das características da informação em suporte digital, especialmente o de fato de que, sempre que são compartilhadas, e, muitas vezes,

pela sua mera utilização ou armazenamento, são copiadas, ou seja, são facilmente difundidas para diferentes dispositivos. Por outro lado, a utilização de dados pessoais em sistemas de inteligência artificial guarda contornos ainda mais desafiadores, podendo-se argumentar ser impossível a "exclusão" dos dados do usuário sem uma interferência extrema em todo o sistema neles (apenas parcialmente e provavelmente de forma mínima) baseado. Desta forma, nas ocorrências mais simples, é necessário que toda a cadeia de agentes com os quais os dados tenham eventualmente sido compartilhados sejam acionados, e sejam todos capazes de identificar os usos secundários do dado desde a entrada em seu poder, e eliminar a sua ocorrência.

Neste sentido, e de forma pragmática, parece que o direito de requerer exclusão de dados pessoais deverá incidir sobre aspectos específicos, i.e., solicitar a exclusão de dados contidos em determinado banco de dados. Na verdade, aparentemente estamos, nestas situações – caso estejamos falando de ações factíveis – a falar de um controle *sobre determinado uso dos dados pessoais*. Ou seja, o direito à exclusão poderá assegurar a um *titular* requerer que determinado dado pessoal mantido para uma determinada finalidade seja excluído daquele banco de dados, e ainda que sejam excluídas suas cópias ocorridas no âmbito das utilizações posteriores que dele decorrem. No entanto, uma vez que a "informação", embora conceitue um objeto de forma relativamente aberta, deve finalmente se referir a um objeto existente no mundo para sobre ele atuar (HALLINAN e GELLERT, 2020), é preciso que a prescrição legal se direcione a esta materialidade para efetivar suas prescrições.

3.3 Relacionando-se por meio de interfaces: aplicação de teorias da argumentação multimodal ao conceito de *privacy by design*

Nesta seção, avaliamos a aplicabilidade de uma teoria multimodal discursiva, conforme formulada por Bateman & Wildfeuer (2014), ao estudo da tomada de decisão relacionada a dados pessoais em interações mediadas por interfaces de aparelhos eletrônicos. Uma teoria acerca dos sinais, sua percepção e efeitos seletivos de Innes (2004) contribuirá para o arcabouço teórico de análise. Propõe-se ainda o desenvolvimento e adoção de métrica baseada no critério de articulabilidade de Groarke & Palczewiski (2016) para mensurar a capacidade de determinação, dependente do contexto e do sujeito, dos elementos discursivos e perceptivos capazes de atuar sobre o plexo de influências sobre o indivíduo até o momento da tomada de decisão, cuja catalogação, estruturação e análise realizar-se-ão conforme o referido modelo

desenvolvido por Bateman & Wildfeuer (2014), considerando a presença de atributos discursivos nas sequências visuais que usualmente enquadram a interação de pessoas com aparelhos eletrônicos, inclusive no contexto de tomada de decisões em que em algum grau são disponibilizados dados pessoais a terceiros, frequentemente, mas não apenas, no intuito de obter acesso a serviços.

3.3.1 Sinais: seletividade e impactos no estabelecimento do contexto interpretativo possível

Ao avaliar sobre riscos à privacidade, as pessoas não são capazes de apreender todos os riscos e benefícios embutidos – que muitas vezes sequer poderiam ser mensurados mesmo por um observador com posição privilegiada –, tampouco decidem uniformemente ao longo do tempo. Essas decisões "dependem de uma miríade de fatores: sobre o que estão pensando no momento, quanto tempo levam para tomar uma decisão, o quanto estão cientes de potenciais riscos à privacidade, e assim por diante" (SOLOVE, 2020). Como observaram Acquisti, Taylor e Wagman (2016), pequenas mudanças de contexto e cenários podem conduzir a conclusões completamente divergentes relativamente à tendência de consumidores a pagar para proteger seus dados pessoais.

Assim, antes de passar à avaliação de parâmetros para avaliar as situações específicas em que sujeitos são confrontados por situações envolvendo proteção de dados pessoais e instados a tomar decisões específicas acerca do risco-retorno que representam, cabe verificar o contexto no qual se insere tal tomada de decisão no tocante à percepção genérica de risco, ou seja, quais seriam os parâmetros para compreensão individual de que determinada situação ou ambiente possa ser considerado perigoso *a priori*, uma vez que "indivíduos, ao navegarem pelo espaço social, estão constante e ritualisticamente buscando por 'sinais de alerta' e 'sinais para alerta'" (GOFFMAN, 1972). Formula-se a hipótese de que a ausência de sinais de alerta na maior parte das situações em que se toma decisões relacionadas ao risco-retorno da disponibilização de dados pessoais possa contribuir para uma preponderância da consideração dos benefícios imediatamente percebidos como advindos de tal disponibilização, levando a uma ponderação falha, uma vez que faltariam parâmetros para sopesá-los com potenciais externalidades negativas de tal conduta.

Por outro lado, é provável que um desconhecimento acerca das externalidades positivas obtidas pelos agentes de tratamento de dados pessoais a partir de tais atividades seja responsável por um enfraquecimento da posição individual quanto à eventual poder de barganha em trocas que envolvam ceder dados pessoais para obtenção de produtos ou serviços (MALGIERI e CUSTERS, 2017). Mais do que isso, os indivíduos não são capazes de *visualizar* a lógica por detrás dos serviços que utilizam e de como eles geram informações relevantes para atividades essencialmente *descorrelacionadas* a estes serviços, sendo-lhes oferecidas apenas "intangíveis e inquantificáveis transações com seus dados", o que "sistemática e significantemente dificulta o surgimento de arranjos econômicos mais favoráveis" (HAYNES e CAROLYN NGUYEN, 2014). Isso se reflete em incapacidade de sequer *vislumbrar* o que se pode sofrer ou ganhar a partir da cessão de dados pessoais.

Nesse sentido, considerando que o caráter arquitetônico dos ambientes virtuais é estabelecido pela literatura, propõe-se abordar as percepções individuais genéricas sobre risco quando da sua interação com ambientes virtuais a partir dos cânones propostos por Innes (2004), em que se propõe que "pessoas interpretam a ocorrência de determinados incidentes como 'sinais de alerta' sobre os níveis de risco a que estão sujeitos ou potencialmente expostos", constituindo tais sinais importantes referências para construção simbólica do espaço social.

Construindo sobre as bases teóricas de Goffman (1972) e Eco (1976), Innes (2004) define um sinal como um signo que faz algo, ou seja, que opera um efeito. Esta conexão de um efeito a um signo depende, por sua vez, de convenções sociais, que são contextualmente situadas. Assim, teoriza que um sinal é composto de três elementos, a saber: uma expressão, ou seja, sua descrição denotativa; um conteúdo, ou seja, a conotação atribuída; e um efeito, que corresponde à uma possível modificação no comportamento em decorrência da ocorrência do sinal. A junção desses três componentes diferencia sinais de meros ruídos, estes últimos definidos como o plexo de incidentes e ocorrências que não assumem significância real para as pessoas e suas rotinas (INNES, 2004).

Ao considerar o papel do risco na funcionalidade comunicativa dos sinais, Slovic (1992; 2000) procurou responder, sob a ótica da psicologia social da percepção de risco, o porquê de pessoas voltarem sua atenção mais a alguns riscos do que a outros, especialmente diante da observação de análises epidemiológicas cuja conclusão remete ao fato de que os riscos de que as pessoas mais sentem medo tendem a não ser aqueles que representam as ameaças mais objetivas. Sua análise sugere que diferentes riscos tendem a ter diferentes "valores de sinal"

para suas diferentes audiências sociais (SLOVIC, 1992; 2000). Innes (2004) teorizou ainda acerca de "sinais fracos" que, vistos isoladamente, não produzem efeitos – ou seja, não se comportam como sinais – , porém, diante de uma "exposição cumulativa a uma série de sinais fracos pode ser interpretado como um 'sinal forte'", o que chamou de efeito amplificador, destacando que sua pesquisa empírica indicou que a proximidade pode ser temporal ou espacial, bem como que a interferência na interpretação dos signos pode se dar com relação a signos percebidos anteriormente ou que venham a ser percebidos no futuro, ou seja, prospectiva ou retrospectivamente (INNES, 2004, p. 12). Estes achados são consistentes com aqueles de estudos recentes com relação ao papel da percepção sobre proteção de dados e interesse próprio como *drivers* de aceitabilidade moral quando pessoas são instadas a ceder dados pessoais (KODAPANAKKAL et. al, 2020).

A pesquisa empírica realizada por Innes (2004) o levou a teorizar ainda acerca da noção de "sinais de controle", que são gerados por agências de controle social e que poderiam trazer senso de segurança ou insegurança, a depender do contexto em que se inserem. Além disso, Innes (2004) observou que os *efeitos* dos sinais não se limitavam ao medo, mas frequentemente a outras emoções, como raiva, preocupação, sentimento de segurança, vulnerabilidade etc. Genericamente, argumentou que os sinais podem gerar efeitos "afetivos", alterando a forma como pessoas se sentem; "cognitivos", alterando a sua maneira de pensar; ou "comportamentais", alterando os mecanismos de ação individual. Não é difícil enxergar a falácia da autodeterminação como um *sinal de controle* que reforça nos indivíduos a sensação de segurança ao adentrar ambientes digitais e os faz deixar de se comunicar adequadamente, e.g., buscar ativamente por sinais de controle e de alerta (INNES, 2004), negociar a visibilização e invisibilização de práticas sociais (LYON, 2017) e informações.

Por outro lado, Innes (2004) observou que narrativas mediadas, bem como imagens, exercem influência potente na comunicação acerca do medo, e que sinais mediados frequentemente são responsáveis por uma função de 'enquadramento' para indivíduos em termos de como eles interpretam e definem seus encontros e experiências. Chama atenção para o fato de que o foco não é nas qualidades inerentes a um incidente em si, mas na forma com que este é interpretado e construído simbólica e significativamente por indivíduos, comunidades e audiências, e que a percepção de risco não é fenômeno individualizado, mas socialmente fundado e moldado (INNES, 2004).

Sintetizando esta seção para a nossa discussão, inicialmente parece que a existência de sinais de alerta com relação ao compartilhamento de dados pessoais poderia desencadear uma maior atenção com relação a estes riscos. Isto certamente não resolveria as dificuldades relacionadas à questão, mas poderia ir na direção de promover um maior equilíbrio nas condições de ponderação do sujeito nas situações relativas a dados pessoais que dependem de tomada de decisão individual. Por outro lado, o estudo de Innes também identificou "*control signals*" que, a depender do contexto, tendiam a deixar a população mais tranquila com relação à inexistência de riscos em um dado ambiente, bem como que sujeitos tenderam a enquadrar as suas concepções acerca do espaço social a partir de determinados sinais a que se atribuiu importância. Considerando a preponderância de engenharia de gamificação nos aparelhos eletrônicos que usualmente medeiam trocas "voluntárias" de dados pessoais, pode ser que os sinais disponibilizados ao usuário ao longo da experiência que levará à coleta de dados pessoais contribuam para que a percepção de risco não seja preponderante no processo de tomada de decisão relacionado a dados pessoais, mas, ao contrário, que contribuam para prevenir que adentrem à esfera de consciência as externalidades negativas decorrentes da cessão de dados pessoais.

No próximo tópico, a partir de teorias acerca da multimodalidade discursiva, buscar-se-á compreender de forma mais específica a interação entre sinais transmitidos a sujeitos no momento em que são instados a tomar decisões relacionadas a seus dados pessoais, especialmente com relação à coerência entre si do conteúdo das informações apresentadas ao sujeito sobre a atividade de tratamento de dados a ser desempenhada e suas externalidades, a partir da potencial identificação dos elementos discursivos e persuasivos presentes no momento da interação.

3.3.2 Interfaces e argumentos visuais

Uma interface não é produto de uma máquina, mas de pessoas: assim, dos dois lados da interação com um aparelho eletrônico existem pessoas, e, com isso, *intencionalidade*. Teorias acerca da argumentação visual e multimodal permitem apreender este fenômeno comunicativo e *pessoalizar* os agentes de tratamento de dados em sua interação com usuários de serviços digitais.

Em seu estudo sobre argumentos visuais, Groarke & Palczewiski (2016) argumentam que a relevância da disciplina é inerente no tempo em que os desenvolvimentos tecnológicos tornaram o visual um elemento pervasivo de comunicação, constatando que "as pessoas hoje vivem em uma realidade que não é meramente visualmente permeada, mas visualmente mediada" (GROARKE e PALCZEWISKI, 2016, p. 233). Conclui que "se a teoria da argumentação for incapaz de auxiliar as pessoas a navegar estes espaços pela deliberada, estratégica e justificada utilização de imagens e de sua resposta a elas, tornar-se-á crescentemente obsoleta nas vidas de pessoas digitalmente conectadas" (GROARKE e PALCZEWISKI, 2016, p. 233). Para Kjeldsen (2015), o significado de qualquer argumento deve ser compreendido como uma função da interação entre o argumentador e a audiência em que se insere, e Groarke & Palczewiski (2016) ressaltam a função da "contestabilidade como atributo inerente da argumentação visual" (GROARKE e PALCZEWISKI, 2016, pp. 220-1).

Ainda sobre o caráter argumentativo de imagens, elementos ou sequências visuais, Groarke & Palczewiski (2016) argumentam que a importância da nomenclatura *argumentos visuais* reside no reconhecimento de que há formas diversas de comunicação, que devem ser recebidas ou estudadas diferentemente. Reconhecer argumentos visuais é uma forma de reconhecer que isso pode ser feito pela produção e apresentação de imagens de várias formas, não apenas pela produção de palavras e frases (GROARKE e PALCZEWISKI, 2016, p. 225), enquanto a compreensão e avaliação de argumentos visuais depende do desenvolvimento de habilidades específicas, o que foi chamado de literacia visual (LANGSDORF, 1996), ou seja, a capacidade de *articular* as influências atuantes sobre o sujeito mediante a ação de imagens.

Groarke & Palczewiski (2016) destacam, no contexto da importância de desenvolver algo como uma *literacia visual*, a essencialidade de se "reconhecer quando instrumentos de argumentação ou persuasão, sejam visuais ou verbais, são projetados, empregados ou nos afetam de formas que subvertem, perpassam ou sobrepõem uma resposta racional a ele" (GROARKE e PALCZEWISKI, 2016, p. 230). Disso decorre que a capacidade de identificar os contextos em que tais mecanismos "a-racionais" – ou seja, não decorrentes de um raciocínio logicamente estruturado – estejam em cena pode tornar os sujeitos sensíveis à percepção de que devem buscar motivações racionais antes de tomar a decisão a que são instados – articulando, portanto, estes mecanismos "a-racionais". Em outras palavras, "o reconhecimento de um agente das forças persuasivas agindo sobre ele afigura-se como precondição essencial para o seu engajamento crítico com esses instrumentos e agentes de persuasão" (GROARKE e PALCZEWISKI, 2016, p. 232). Como parâmetro para verificação das condições de

possibilidade ou da ocorrência desse fenômeno, que chamou de monitoramento reflexivo, Groarke & Palczewiski (2016) propuseram o critério da *articulabilidade*, isto é, a mensuração do grau com que o indivíduo é capaz de articular as forças lógicas e persuasivas que estejam atuando em um dado momento, o que serve a objetivos analíticos e normativos. Analíticos, pois, torna-se possível estruturar conhecimento em torno dessas influências; normativo, pois, o sujeito torna-se capaz de agir a partir de elementos persuasivos racionalmente estruturados, normatizando a sua conduta conforme estes critérios.

A adoção de um modelo argumentativo capaz de conduzir à avaliação conjunta dos elementos visuais e verbais que medeiam as relações de sujeitos com agentes de tratamento de dados parece ter o potencial de melhorar consideravelmente a capacidade de avaliação das condições de articulabilidade presentes nas interfaces por meio das quais os usuários são instados a tomar decisões acerca da disponibilização de seus dados pessoais, viabilizando a superação da dicotomia decorrente da equivocada noção do *paradoxo da privacidade* Solove (2020).

3.3.3 Interfaces e discursividade multimodal

A partir da compreensão de que, ao navegar por ambientes virtuais, os indivíduos são frequentemente confrontados com situações em que são instados a tomar decisões relativas a seus dados pessoais, e.g., a disponibilização de determinado dado pessoal, para determinado agente, para determinada finalidade, e que tal *solicitação à ação* ocorre *por meio de* dispositivos eletrônicos que apresentam informações de diversas formas, i.e., imagens, texto, sons, vibrações etc., parece valioso o emprego de um modelo capaz de estruturar as informações transmitidas pelas interfaces de dispositivos eletrônicos e compreender a forma com que dialogam com o agente instado a decidir a partir da contribuição de cada elemento para este processo que se propõe avaliar sob a ótica *discursiva*.

Assim, de forma a viabilizar estudos empíricos que expliquem aspectos da tomada de decisão acerca de *risco-retorno* no momento da disponibilização de dados pessoais que superem a falácia do paradoxo da privacidade (SOLOVE, 2020), propõe-se a adoção de modelo teórico de discurso multimodal conforme estruturada por Bateman & Wildfeuer (2014). O frame*work* desenvolvido pelos autores assume que as tarefas em que estão envolvidos os participantes de um discurso multimodal não devem ser descritas em termos de resolução geral de problemas,

mas, ao contrário, devem sê-lo em termos de um desenvolvimento mais focado nas funções da linguagem – de compreensão de quais propósitos discursivos estão sendo objetivados. Partindo do pressuposto que as interfaces de dispositivos eletrônicos representam sequências visuais criadas para empreender atividade comunicativa com o usuário de tais dispositivos e guiar a atividade interpretativa com relação à sua funcionalidade e potencialidades, o modelo de Bateman & Wildfeuer (2014) mostra-se compatível e potencialmente útil no estudo de ocorrências no contexto de interações de pessoas com dispositivos eletrônicos dotados de *tela*, ou, de forma mais genérica, *interface* de comunicação.

Os autores constroem o que chamam de "teoria multimodal do discurso" pragmaticamente fundada, pela característica de "endereçar o relacionamento entre alguns artefatos comunicativos e o contexto de seu uso, enquanto mantem a relação próxima com os detalhes técnicos de um artefato como provendo parâmetros ou instruções para levar a cabo aquela contextualização" (BATEMAN e WILDFEUER, 2014, p. 185). Para tanto, partem da definição de modos semióticos, argumentando que "a semântica do discurso provê os mecanismos pragmáticos interpretativos necessários para relacionar as formas com que um modo semiótico se distingue de seus contextos de uso, e para demarcar o *escopo intendido* de interpretação dessas formas" (BATEMAN e WILDFEUER, 2014, p. 183). Esta abordagem procedimental, ou dinâmica, da interpretação, é baseada em estudos acerca do processamento cognitivo do discurso que propõem que "a compreensão opera pela construção explícita de modelos mentais representativos de situações que um texto descreve" (BATEMAN e WILDFEUER, 2014, p. 184). Parte-se de estudos que atribuem característica de 'multidimensionalidade' a tais modelos, que se constituem, ao menos, de aspectos espaciais, causais, temporais e motivacionais, que operam por meio de procedimentos de atualização do discurso (BATEMAN e WILDFEUER, 2014, p. 184). O modelo proposto emprega o discurso semântico dinâmico com a finalidade de prover uma caracterização abstrata, porém formalmente detalhada, de como o processo de construção de significado é guiado pela informação linguística presente no momento da interação, conjuntamente com outros conhecimentos semânticos e contextuais, sendo capaz de "capturar formalmente a ideia de que o processo interpretativo do discurso ocorre pela progressiva adição de informações a um contexto discursivo crescente" (BATEMAN e WILDFEUER, 2014, pp. 184-5), especificando-se formalmente a semântica em termos de *princípios de atualização semântica do discurso*, que dispensam a necessidade de se determinar *a priori* as possíveis capacidades-significativas, ou na linguagem de Gibson (1979), *affordances*, embora o processo ultimamente os torne

cognoscíveis. O modelo parte, com isso, de um modelo de modos semióticos que combina formal e detalhadamente informações acerca do componente discursivo, considerações sobre a materialidade e sobre a percepção desse material (BATEMAN e WILDFEUER, 2014, p. 186). Em termos simples, torna-se possível identificar a forma com que cada elemento de informação apresentado ao sujeito é percebido e processado, em conjunto com outras informações recebidas e com outras de que dispusera o sujeito, e por ele articulado (GROARKE E PALCZEWISKI, 2016).

Sob o ponto de vista da adequação metodológica e temática do modelo, consistente com o tipo de problema que se endereça, observa-se que a *Segmented Discourse Representation Theory* (SDRT) (ASHER e LASCARIDES, 2003) é uma abordagem que viabiliza a caracterização precisa dos aspectos dinâmicos do processo descrito, que se chamou de "interpretação guiada por artefato" (BATEMAN e WILDFEUER, 2014, p. 186), e sobre a qual foi construído o modelo de Bateman & Wildfeuer (2014) cuja empregabilidade no contexto analisado se propõe. Um aspecto distintivo do *framework* SDRT que se destaca é a sua base em "lógicas distintas que se combinam para formar uma 'lógica geral da interpretação do discurso'" (BATEMAN e WILDFEUER, 2014, p. 186), que resulta em atributo de modularidade que permite lidar com questões complexas por meio de operações relativamente simples (BATEMAN e WILDFEUER, 2014, p. 186). A interpretação do discurso no SDRT opera "pela construção de uma representação semântica para cada contribuição entrante para o discurso [...], que é então conectada por meio de relações discursivas em uma estrutura discursiva crescente" (BATEMAN e WILDFEUER, 2014, p. 187). As relações discursivas são definidas de modo a explicitar tanto a sua aplicabilidade a representações semânticas específicas, quanto as demandas que faz com relação ao contexto. Então, "olha-se tanto no sentido de cima para baixo, no sentido de formas linguísticas concretas e sua semântica, quanto no sentido de baixo para cima, no sentido do contexto" (BATEMAN e WILDFEUER, 2014, p. 187). As demandas da relação discursiva "definem precisamente as formas com que 'gaps' identificáveis na interpretação são tanto criados, quanto resolvidos" (BATEMAN e WILDFEUER, 2014, p. 187).

No modelo de Bateman & Wildfeuer (2014), ora proposto, a forma geral deste *framework* é utilizada para caracterizar o estrato semântico do discurso dos modos semióticos, independentemente de sua materialidade. Neste sentido, é hábil a servir para a análise das diversas formas com que a interação discursiva com aparelhos embutidos de tecnologias da informação ocorre. Destaca-se que as relações discursivas são caracterizadas por duas

perspectivas distintas, a saber: a primeira se refere a "limitações duras" que precisam ser respeitadas pelo conhecimento do contexto para que se obtenha a relação discursiva, e que são chamados *postulados de significado* que seguem a seguinte regra geral: $\varphi_{R(\alpha,\beta)} => conditions(\alpha,\beta)$[7] (1) (ASHER e LASCARIDES, 2003). A segunda perspectiva provê regras de inferência abdutivas (em contraste às anteriormente descritas, tidas por *não*-abdutivas em sua expressão monotônica), chamadas *axiomas padrão*, "que especificam que relações discursivas podem ser aplicáveis dados atributos especificados dos elementos discursivos sendo relacionados" (BATEMAN e WILDFEUER, 2014, p. 187). O esquema desta regra é descrito da seguinte forma: $(?(\alpha,\beta,\gamma)\,\Delta\, some\ stuff > R(\alpha,\beta,\gamma)$ (2) (ASHER e LASCARIDES, 2003), em que $?(\alpha,\beta,\gamma)$ "indica uma relação discursiva subespecificada entre os segmentos α e β do contexto da estrutura discursiva etiquetada de γ, R é a relação discursiva abduzida específica, e > é a implicação possível, tipicamente lida como 'se...então usualmente...'", enquanto *some stuff* é utilizado "para representar as condições que devem ser mantidas na ordem antecedente para que haja evidências em favor da relação (BATEMAN e WILDFEUER, 2014, pp. 187-8), i.e., para que se possa racionalmente afirmar a probabilidade de sua ocorrência.

A importância dessa separação de perspectivas entre hipóteses abdutivas e consequências necessárias está em prover uma conexão manejável entre descrições do mundo e conhecimentos prévios externos ao texto, ou seja, permite compreender como os conhecimentos e preferências presentes na "bagagem pessoal" do sujeito, inclusive preferências sobre privacidade, se relacionam com as influências exercidas pelos elementos (i.e., argumentos, sinais e ruídos) que lhe são apresentados na interação com a interface e aqueles que evocam.

> Os axiomas padrão descrevem que informações o contexto discursivo precisa prover para interpretar as relações discursivas entre os segmentos, o que então oferece os "mecanismos necessários para identificar '*gaps'* determinados textualmente que precisam ser preenchidos a partir do contexto para que um discurso coerente resulte (BATEMAN e WILDFEUER, 2014, p. 188).

Neste ponto, quer-se dizer que, na interação com a interface, o indivíduo é frequentemente instado a agir, e, em cada um destes momentos, terá à sua disposição os referidos elementos e um determinado contexto. Desta forma, o processo argumentativo como

[7] "The lefthand side of the rule picks out a particular discourse relation R being added to the current discourse structure between the segments labelled a and b. If this relation holds, the conditions on the righthand side are required to follow by regular, non-defeasible material implication". (Bateman & Wildfeuer, 2014, p. 187).

um todo é "descrito como uma 'colagem' das formas lógicas de clausulas de acordo com axiomas padrão aplicáveis para construir uma forma lógica geral com o 'máximo de coerência' para um dado discurso" (BATEMAN e WILDFEUER, 2014, p. 188). Assim, e sintetizando o exposto, o SDRT formula o processo de "'colagem' semântica dos elementos discursivos em termos de um procedimento de atualização do discurso" (ASHER e LASCARIDES, 2003), construindo estruturas progressivamente maiores pela invocação de relações discursivas que impõem segmentações sobe coleções de estruturas representativas de discurso introduzidas por um discurso, operando-se a seleção de relações discursivas sempre no sentido de tentar 'maximizar' a coerência do discurso como um todo, o que pode levar à aceitação ou rejeição de hipóteses anteriormente construídas a partir da disponibilização de novas informações (BATEMAN e WILDFEUER, 2014, p. 189). Vê-se que é um modelo adequado para uma investigação fenomenológica e para a construção de teorias enraizadas (GLASER et al., 1967) acerca da forma com que pessoas que não dispõem de conhecimento específico acerca do funcionamento de computadores lidam com estes sistemas, e, assim, identificar as situações em que estão expostas à criação e exploração de vulnerabilidades (CALO, 2017), e em que o direito deverá eventualmente agir *ex ante*.

O modelo de Bateman & Wildfeuer (2014) estende o *framework* do SDRT, estabelecido para o discurso verbal, para o discurso multimodal em geral. Para tanto, são incorporados os diferentes papeis da percepção quando se lida com artefatos não-linguísticos, de forma a evitar consequências comuns de sua desconsideração, especialmente que se lide com processos de construção de conhecimento, notadamente aqueles decorrentes da interação com artefatos não-linguísticos, como meramente convencionais, desconsiderando importantes fontes de significado (BATEMAN e WILDFEUER, 2014, p. 189). Neste ponto, remete-se ao estudo de Innes (2004) como forma de destacar o papel que determinados sinais podem ter na construção de significado acerca do espaço social e na geração de efeitos nos agentes, ressaltando-se a alta seletividade destes sinais, ou seja, a característica de que enquanto alguns aspectos potencialmente percebíveis no ambiente não recebem relevante consideração no processo de atribuição de significado, outros exercem efeitos diversos no indivíduo, sem que se possa necessariamente esclarecer em termos de critérios estritamente lógico-racionais definidos *a priori* as razões para esta distinção. Isto não deve representar, outrossim, que o processo seja *incompreensível*.

Os autores propõem que alguns dos 'gaps' que demandam preenchimento para a consecução de uma estrutura discursiva coerente operam também como *drivers* de objetivos e

hipóteses durante a percepção visual, e que o "material visual 'perceptivamente interpretado' estabelece uma alternativa apropriada às representações lógicas trazidas pela semântica compositiva no caso de interpretação linguística do discurso". Em outras palavras, a "percepção revela unidades analíticas potenciais, que a organização discursiva então refina e seleciona quando coloca em uma estrutura comunicativamente motivada", o que provê a "matéria-prima para a formação de hipóteses discursivas em mídias visuais" (BATEMAN e WILDFEUER, 2014, pp. 189-90). A partir da forma lógica como um todo obtém-se material para o processo de "maximização abdutiva da coerência discursiva pela busca por relações discursivas aplicáveis". Em contraste ao papel rígido da forma lógica no modelo tradicionalmente aplicado apenas ao discurso verbal, aqui "as representações já são o resultado de extensiva racionalização abdutiva, de um lado, e de informações 'diretamente' percebidas, de outro (BATEMAN e WILDFEUER, 2014, pp. 191-3). É precisamente esta relação entre formação de hipóteses discursivas e percepção que sugere estratégias potenciais de aplicação em investigações empíricas, e pode viabilizar, por exemplo, o esclarecimento da forma com que um contexto discursivo dinamicamente desenvolvido pode contribuir para o isolamento de significados intendidos por signos visuais específicos (BATEMAN e WILDFEUER, 2014, p. 193), permitindo estabelecer os elementos que agem sobre o agente instado a decidir, inclusive aqueles a- ou extra racionais, e ainda compreendendo se são internos ou externos ao argumento, permitindo cumprir com o critério proposto de articulabilidade (GROARKE e PALCZEWISKI, 2016).

O modelo, conforme destacam seus formuladores, é compatível com modelos de percepção visual que "tratem a percepção como um processo de interrogação ativa do ambiente para a obtenção de informação útil, e não apenas consumo passivo", ressaltando que é bem estabelecido na literatura o fato de que esta "interrogação é altamente sensível tanto a objetivos quanto a tarefas específicos sendo levadas a cabo pelo agente percebedor, e altamente seletiva", apresentando-se o modelo explícito de percepção em termos de *ganho informacional,* por meio do qual a atenção é direcionada às partes do campo perceptivo que proverão máxima informação para resolver as hipóteses relativas ao que está sendo percebido. Neste ponto, ressalta-se aspecto delineado anteriormente neste trabalho: as formas com que indivíduos interagem – direta ou indiretamente – com sensores, e desta forma geram informações as mais diversas, no mais das vezes não se relaciona estreitamente com a atividade que origina a coleta, que, por sua vez, não está correlacionada com as atividades eventualmente desenvolvidas a partir destas informações, i.e., a distribuição de suas ocorrências frequentemente se sobrepõem

e se afetam mutuamente, mas não guardam relação de causalidade, e, com isso, atuar sobre uma não exerce necessária (nem provavelmente) os efeitos desejados. Ademais, ao manter estas ocorrências atadas, inclusive a partir de expressos princípios como o da finalidade específica das atividades com dados pessoais e de posição legal de destaque ao consentimento e outras formas de controle individual sobre a informação, mantém-se distantes da maior parte da sociedade – "titulares" – as atividades realmente significativas desenvolvidas a partir dos dados gerados pelos sensores ambientais e embutidos em dispositivos com que interagem diuturnamente boa parte dos cidadãos do mundo. As possibilidades de coleta e processamento de dados, por outro lado, são inúmeras e rotineiramente desenvolvidas e adotadas pela sociedade; a forma com que o direito se relaciona com estas possibilidades e é capaz de agir sobre vulnerabilidades (CALO, 2017) é determinante para a sua relevância e capacidade de atingir objetivos caros à sociedade, devendo haver uma consonância entre estes papeis do conceito de informação que atraem aplicabilidade da norma, e aquele referente ao meio sobre o qual se deve agir (HALLINAN e GELLERT, 2020).

No contexto de uma vida social em que o papel da mediação da tomada de decisões relativas a dados pessoais (e muitas outras) é realizado por interfaces que exibem sequências visuais ininterruptamente, a caracterização dessa interação em termos *discursivos* e *perceptivos,* e sobretudo no *framework* da teoria de discurso multimodal delineada por Bateman & Wildfeuer, pode permitir que se alcance o requisito de articulabilidade proposto por Groarke & Palczewiski (2016) para determinar os elementos racionais e a-racionais que atuam sobre o sujeito que toma parte nesta relação, indo além das soluções tradicionalmente oferecidas ao se abordar a proteção de dados pessoais prementemente pela ótica do direito à privacidade, e sobretudo a partir da ótica do paradoxo da privacidade, que conduzem frequentemente ora à conclusão de desvalorização da privacidade por parte do cidadão médio, ora à da completa ausência de condições de manifestação de vontade dado o notório projeto de tais interfaces de forma a conduzir a experiência do usuário (SOLOVE, 2020). Estas abordagens não têm levado a soluções satisfatórias, seja no sentido de proteger titulares de dados, seja no de viabilizar atividades significativas para a sociedade a partir de tratamento de dados pessoais.

Propõe-se a aplicação do modelo apresentado para a condução de estudos empíricos destinados a avaliar sequencias, visuais e apresentadas de outras formas, em que são tomadas decisões relacionadas a dados pessoais, e.g., situações em que se concede acesso a dados pessoais para acesso a determinado serviço. O objetivo seria viabilizar a identificação dos elementos discursivos e perceptivos atuantes sobre o agente e a forma com que se adicionam à

estrutura discursiva até a identificação e preenchimento de ‘gaps’ interpretativos pelo recipiente, i.e., a identificação dos riscos e benefícios envolvidos; das informações disponíveis, disponibilizadas ou obtidas previamente que são solicitadas a tomar parte no processo decisório; das razões determinantes para a tomada de decisão; a tomada de decisão em si. Propõe-se o estabelecimento de parâmetro quantitativo-qualitativo baseada no critério de articulabilidade de Groarke & Palczewiski (2016) para mensurar as condições que determinados sujeitos, em determinadas situações, detém para o exercício desta tomada de decisão, o que pode conduzir a descobertas relevantes para aprimorar a regulação relativa a proteção de dados pessoais, por exemplo, a partir do emprego de *nudges* (SUNSTEIN, 2009) que favoreçam escolhas condizentes com os valores perseguidos pela sociedade e pelo direito.

4 MERCADOS DE DADOS E CONCORRÊNCIA

Neste capítulo, trata-se, inicialmente, de relações entre disposições e categorias do GDPR, de um lado, e características da economia contemporânea, de outro, observando eventuais adequações ou incongruências. Em seguida, são expostos e avaliados resultados de estudo empírico a respeito do resultado do GDPR sobre os mercados que se utilizam de dados pessoais.

4.1 GDPR e *big data*: Incompatibilidade?

A relação básica do direito com as atividades realizadas com *big data* pode ser compreendida pela forma com que se descreve usualmente o conceito: a partir dos “quatro Vs”, relativos ao volume dos dados coletados, à variedade de suas fontes, à velocidade com que a análise pode se desenvolver, e a veracidade da informação. Zarsky (2016) adota em sua análise uma definição ampla do conceito, voltando-se a apreender a

> maneira fundamental com que dados são coletados, armazenados e subsequentemente utilizados – todos como resultado de recentes desenvolvimentos tecnológicos. Na atual era digital, dados são coletados utilizando múltiplos sensores bem como diversas aplicações que gravam movimentos, comunicações e transações dos usuários. Eles são armazenados com sofisticados mecanismos em bancos de dados distribuídos cujos custos decrescem constantemente. Finalmente, é utilizada em processos analíticos avançados e aplicados em uma miríade de contextos. (ZARSKY, 2016, p. 999)

Além disso, o conceito de *big data* frequentemente se refere a formas específicas de análises de dados, especialmente aquelas realizadas por computadores a partir de mecanismos de mineração de dados, e que permitem que tais análises façam uso de enormes quantidades de dados que estejam à sua disposição, especialmente em contextos em que não se tem um ponto de partida para a construção de conhecimento a partir das informações disponíveis (ZARSKY, 2016, p. 999). Em muitos casos, estes bancos de dados contêm informações de natureza pessoal, ou o processamento dos dados leva à geração de informações de natureza pessoal, ou, ainda, geram informações que são utilizadas para a tomada de decisões relativamente a indivíduos ou grupos específicos, desta forma afetando-os diretamente (ZARSKY, 2016, p. 1000).

Edwards e Veale (2017) chamam atenção para o papel de sistemas de aprendizado por máquina na atual *infraestrutura* da sociedade. Atualmente, este tipo de sistema tem sido empregado para a realização e automação de diversas tarefas, como formas variadas de pesquisa, observação por máquina, reconhecimento de voz e outros elementos humanos e espaciais, cruzamento de pontos de dados para identificação de anomalias em diferentes contextos, dentre inumeráveis outros. O que é de relevante compreensão é o papel central que adquiriram tais sistemas em nossa infraestrutura social, e que ultimamente são empregados com todo tipo de finalidade, inclusive algumas econômicas e políticas com grande importância para a sociedade, tornando a sua integridade vital para a própria normalidade social, econômica e democrática. Neste ponto, é útil referirmo-nos à distinção de Edwards e Veale (2017, p. 26) acerca de formas de aprendizado por máquina que incluem supervisão humana, e aqueles que são programados para, a partir de um certo ponto, construir "sozinhos" o aprendizado e o conhecimento possíveis a partir de determinado conjunto de dados. O "aprendizado supervisionado" recebe um vetor de variáveis (como sintomas físicos ou características), e uma etiqueta de "correto" para este vetor (como um diagnóstico médico ou um agregador válido para as características), que é chamada de "verdade fundamental". "O objetivo do aprendizado supervisionado é prever de forma acurada esta verdade fundamental a partir do *input* de variáveis em casos em que apenas este último é conhecido" (EDWARDS e VEALE, 2017, p. 25). Por sua vez, o "aprendizado não-supervisionado" não é supervisionado por uma verdade fundamental, ou seja, "sistemas de aprendizado por máquina tentam inferir estruturas e grupos com base em outras heurísticas, como proximidade", i.e., características comuns ou *proximidade* entre estas características. Nestes casos, em geral se está interessado em observar que características podem estar "próximas" umas das outras com relação a diversos parâmetros possíveis, i.e., todos aqueles que a máquina seja capaz de receber e processar, sem saber

anteriormente ou de forma imediata que agregado de características próximas poderá emergir da análise (EDWARDS e VEALE, 2017, p. 25). Com esta explicação, vemos que o *valor* do conhecimento que pode ser gerado a partir de práticas de aprendizado por máquina, frequentemente, está relacionado ao fato de ele gerar conhecimentos ou eficiências que não teríamos sem o emprego do sistema, e, com isso, não pode ser completamente determinado ou previsto *ex ante*.

Há aqueles que se posicionam no sentido de afirmar a transformatividade das práticas de *big data* e o seu potencial de, além de gerar eficiência e promover o bem-estar, prover à sociedade amplo e valoroso conhecimento que poderia amparar a melhora da qualidade de vida em diversos contextos (ZARSKY, 2016, p. 1000). Por outro lado, este "fascínio" foi progressivamente decrescendo, com críticos afirmando que o termo 'big data' seria sobretudo, um *hype* (ZARSKY, 2016, p. 1001). Estas críticas fiam-se em duas principais razões: a uma, pode-se dizer que a "revolução do *big data*" seria, em verdade, uma "mera" evolução, referindo-se a práticas, benefícios e problemas que foram continuamente desvelados e desenvolvidos na cultura humana, ao ponto em que foram distintamente capturados pela atenção da mídia e do público. Assim, não teria havido um verdadeiro salto tecnológico que requereria uma recalibração de objetivos de políticas públicas e normas jurídicas (ZARSKY, 2016, p. 1001). Zarsky (2016, p. 1001) reconhece pertinência nesta crítica, mas observa que ela tem pouca relevância para discussões em matéria de políticas públicas, pois, independentemente de o processo ser evolucionário ou revolucionário em sua natureza, não pode ser ignorado pelo direito e pela política quanto às mudanças incrementais – porém altamente pervasivas – trazidas pela "nova era digital". Devem prover respostas a estas mudanças, sejam elas parte de uma revolução ou de uma "mera evolução".

Uma segunda crítica afirma que a noção de *big data* se refere a uma "promessa que não pode ser cumprida", pois os seus benefícios existiriam em teoria, mas, na prática, seriam custosos (tanto em termos monetários, quanto em termos de outros custos, como danos causados a titulares de dados) e os seus resultados, assim, inconsistentes (ZARSKY, 2016, p. 1002). No entanto, mesmo que os benefícios propagados das práticas relacionadas ao *big data* sejam frequentemente exagerados, há razões para se partir do pressuposto de que tais práticas *geram benefícios relevantes*, ou apresentam este potencial, especialmente se desenvolvidos de forma harmônica com interesses sociais relevantes. Por outro lado, sob um ponto de vista econômico, inovações tecnológicas pertencem à categoria de objetos e serviços referidos como "bens públicos", ou seja, aqueles que ostentam a característica distintiva de poderem ser

replicados com baixo custo marginal e serem "não-rivais", i.e., a sua fruição por um indivíduo não inviabiliza a igual fruição por parte de outros. Em conjunto, estas características representam o risco de que o ritmo de desenvolvimentos tecnológicos seja menor do que o desejável, dada a ausência de incentivos naturais para se engajar neste tipo de atividade (FISHER, 2001).

Esta tensão entre análises de dados e proteção de dados não escapou ao legislador europeu quanto da elaboração do regime jurídico do GDPR. "Os responsáveis pelo GDPR foram chamados a contrabalancear a habilidade de se engajar em atividades de análise de *big data* sem restrições, e a proteção de direitos e interesses relacionados à privacidade" (ZARSKY, 2016, p. 1002). É precisamente a natureza deste balanceamento e suas falhas que Zarsky (2016) se propôs a avaliar. Vale notar que nem todos aquiescem com esta ideia do reconhecimento de uma *tensão* na disciplina europeia de proteção de dados pessoais: há os que argumentam que o regime de proteção de dados em verdade *promoveria* práticas de análises de dados e ampliaria os benefícios delas decorrentes, especialmente a partir do desenvolvimento de uma relação de *confiança* entre agentes de tratamento de dados e titulares. Esta foi a posição externada oficialmente pela União Europeia em alguns documentos (ZARSKY, 2016, p. 1002). Na literatura jurídica brasileira sobre o tema, também foi notada (ZANATTA, 2019). No entanto, a ideia de que maior tutela a dados pessoais, na forma erguida na UE, promoveria práticas de *big data* não guarda relação com o que pode ser atualmente observado na realidade, e provavelmente se relaciona mais a um desejo do que a previsão baseada em informações consistentes (ZARSKY, 2016, p. 1003).

Zarsky (2016) argumenta que "o balanceamento trazido à tona pelo GDPR é inaceitável e subótimo" (ZARSKY, 2016, p. 1004), destacando que, por um lado, o Regulamento mina a possibilidade de se empreender análises de *big data*, enquanto por outro, a ampla disponibilidade de tecnologias de *big data* mina algumas das medidas e das distinções conceituais que o GDPR apresenta. Para demonstrar a pertinência do argumento, Zarsky (2016) analisa os fenômenos a eles inerentes, e o tratamento declinado pelo GDPR, relativamente aos seguintes aspectos-chave: *limitação das finalidades de atividades de tratamento de dados*; *minimização* de dados pessoais; categorias especiais de dados pessoais; e decisões automatizadas. Veremos cada um separadamente, adiante.

4.1.1 Especificação *ex ante* das finalidades

O Article 5(1)(b) do GDPR apresenta a noção fundamental de que dados pessoais devem ser coletados diante de especificação de um propósito explícito e legítimo, e não poderão ser objeto de atividade de tratamento ulterior que seja incompatível com estes propósitos originalmente especificados. A LGPD acolheu esta técnica, e, igualmente, traz no inciso I do seu artigo 6°, que dispõe acerca dos princípios a serem observados por toda atividade de tratamento de dados que intente ser reputada lícita, o princípio da finalidade, que se refere à "realização do tratamento para propósitos legítimos, específicos, explícitos e informados ao titular, sem possibilidade de tratamento posterior de forma incompatível com essas finalidades" (BRASIL, 2018). Note-se que, embora a finalidade específica apareça na LGPD também como elemento caracterizador do consentimento do titular enquanto base legal autorizadora de atividade de tratamento de dados pessoais, conforme definição do inciso XII do artigo 5° da LGPD, a necessidade de que as atividades de tratamento de dados pessoais possuam finalidade especificada previamente não se limita àquelas que tenham como base legal o consentimento, sendo imposta pelo diploma a qualquer atividade de tratamento de dados pessoais que atraia a sua incidência.

São frequentes as avaliações relativamente à incompatibilidade da noção de especificidade das finalidades de atividades de tratamento de dados e as perspectivas quanto às práticas relacionadas a análises de *big data*, que frequentemente envolvem métodos e formas de uso de dados em que não se detêm no momento da coleta uma ideia clara acerca do resultado do processo ou do valor que pode ser gerado a partir dele (ZARSKY, 2016, pp. 1005-6). A adequação à norma da finalidade específica requer que agentes que realizem análise de *big data* informem aos titulares de dados sobre formas de tratamento de dados que realizarão *no futuro*, e ainda que monitorem continuamente, de perto, estas atividades para assegurar que não desbordem consideravelmente desta finalidade destacada inicialmente em qualquer momento do seu ciclo de vida (ZARSKY, 2016, p. 106). Edwards e Veale (2017, p. 24) notam que sistemas algorítmicos são frequentemente empregados com a finalidade de antecipar ocorrências que não são conhecidas, ou para detectar e subjetivamente classificar algo desconhecido, porém cognoscível *de alguma maneira*, utilizando métodos inferenciais ao invés de medições diretas.

No âmbito do direito comunitário europeu, é interessante notar que este conceito, para além de constituir uma das pedras fundamentais do GDPR, consta expressamente do Article

8(2) do *Charter of Fundamental Rights of the European Union*. Desta forma, a sua adoção pelo legislador no momento da estruturação do Regulamento em matéria de proteção de dados não representa surpresa, pois o contrário poderia resultar na invalidação da norma perante o *European Court of Justice* (ZARSKY, 2016, p. 1006). Para além de aspectos tradicionais e relacionados ao mandato constitucional do legislador europeu, Zarsky (2016) identifica diversas justificativas substantivas para a adoção do princípio da finalidade específica para atividades de tratamento de dados pessoais, mesmo no delineado contexto de *big data*. Por exemplo, o embasamento teórico para que o Regulamento se preocupe em assegurar que controladores de dados respeitem o princípio da finalidade específica da atividade de tratamento permitirá aos titulares de dados o exercício de ao menos algum controle sobre suas informações pessoais – e "o controle é uma justificativa central para a proteção de dados na UE" (ZARSKY, 2016, p. 1006). Outro argumento apontado por Zarsky (2016, p. 1007) é apresentado no conhecido *Recital 29 Working Party*, e afirma que o acolhimento deste princípio levaria a maior grau de confiança entre as partes no âmbito dos ecossistemas de dados, bem como incrementaria a competitividade, pelo alegado potencial do princípio em enfraquecer monopólios atuantes neste mercado e incentivar a entrada e competição de *start-ups* e outros agentes econômicos menores para atuar neste ramo, com benefícios para a coletividade, inclusive quanto aos padrões de proteção à privacidade.

Zarsky (2016, p. 1007) apresenta possíveis respostas a estes argumentos, tanto em um nível teórico, quanto em um nível prático. Sob o ponto de vista teórico, há aqueles que argumentam que, na atual era digital, i.e., diante da adoção massiva de dispositivos eletrônicos e serviços digitais pelo público da forma como são oferecidos, poder-se-ia dizer que os usuários "objetivamente capitularam boa parte de seu controle sobre dados pessoais" (ZARSKY, 2016, p. 1007). Com isso, "a intervenção ativa do estado para prover aos indivíduos direitos que eles não necessariamente demandaram pode remontar a paternalismo e minar a autonomia, e, com isso, a natureza deste direito fundamental deve ser objeto de questionamento" (ZARSKY, 2016, p. 1007). Neste ponto, vale remeter às observações de Sunstein (2009) com relação às diferenças entre um *nudge* e um elemento normativo, e ainda acerca da noção de *paternalismo libertário*.

Sob o ponto de vista instrumental, Zarsky (2016) argumenta que é plenamente possível promover a confiança e limitar as possibilidades de abuso a partir do monitoramento de atividades que utilizam dados pessoais, ao invés de buscar evitar que ocorram. Ademais, deve-se questionar se, e em que grau, o princípio da finalidade é capaz de minar a competição nos

mercados relacionados à utilização de dados pessoais. O pesquisador observou que "ele pode agir como um inibidor de competição, pois limita as possibilidades de que *start-ups* obtenham informação em mercados secundários e utilizem-na para entrar em outros mercados" (ZARSKY, 2016, p. 1007), além de assegurar que "apenas monopólios que já tenham acesso a clientes e seus dados possam permanecer ativos em mercados ricos em dados" (ZARSKY, 2016, p. 1007). Zarsky (2016) previu a possibilidade de ocorrência destes efeitos logo quando da entrada em vigor do GDPR. Nos poucos anos que se seguiram, numerosos estudos, como os de Gal e Aviv (2020) e de Johnson et al. (2020), já foram capazes de demonstrar que, de fato, estes e outros fatores têm exercido papel determinante para que o GDPR exerça efeitos negativos, como reforçar a concentração de mercado e reduzir a geração de externalidades positivas para a sociedade a partir do conhecimento gerado por dados, demonstrando a pertinência de se prosseguir com as avaliações sugeridas por Zarsky (2016) e outros estudiosos com relação ao balanceamento de interesses promovido por disciplinas de proteção de dados pessoais calcadas no princípios fundamentais do GDPR, como é o caso da LGPD. Edwards e Veale (2017), por exemplo, sugerem caminhos para o erguimento de interpretações jurídicas e normas que seriam capazes, ao contrário de concepções de direitos intangíveis, de auxiliar na criação de sistemas de aprendizado por máquina mais úteis e mais "explicáveis", i.e., mais compreensíveis quanto ao que visam atingir e as limitações éticas que observam, levando de uma situação de "escravização *pelo* algoritmo" (EDWARDS e VEALE, 2017) para "escravização *do* algoritmo" (EDWARDS e VEALE, 2018), ou seja, a sua captura pelos objetivos fundamentais da sociedade, em contraposição ao serviço de poucos *arquitetos de escolhas* (SUNSTEIN, 2009) em posição privilegiada para a coleta e processamento de dados pessoais (ZARSKY, 2016).

4.1.2 Minimização dos dados

O princípio da *minimização* dos dados é outra pedra fundamental do regime jurídico do GDPR, e está disposto em seu Article 5(1)(c), que dispõe que, quando da utilização de dados pessoais, estes devem ser "limitados ao que for necessário em reação aos propósitos para os quais são tratados" (ZARSKY 2017, p. 1009). Este princípio tem múltiplas dimensões: se relaciona ao escopo e às categorias de dados a serem inicialmente objeto de coleta com relação à(s) finalidade(s) destacada(s); e se refere à duração limitada do ciclo de vida dos dados objeto de tratamento, i.e., não devem ser mantidos pelo controlador indefinidamente, e em regra

deverão ser excluídos após a realização do uso intendido (ZARSKY, 2016). Na LGPD, este também é um princípio expresso no artigo 6°, especificamente no inciso III, em que recebe o nome de *princípio da necessidade*, e é descrito como "limitação do tratamento ao mínimo necessário para a realização de suas finalidades, com abrangência dos dados pertinentes, proporcionais e não excessivos em relação às finalidades do tratamento de dados" (BRASIL, 2018).

Zarsky (2016, p. 1009) observa que, diferentemente do princípio da finalidade, este não é expresso na *European Union Charter of Fundamental Rights*, o que representa maior grau de flexibilidade para o legislador em relação ao GDPR. As justificativas para a sua adoção são tanto intuitivas, quanto instrumentais, i.e., voltadas a uma finalidade prática (ZARSKY, 2016, p. 1009). Intuitivamente, é notável que, ao adotar o princípio, controladores de dados disporão de menos oportunidades para vulnerabilizar e explorar vulnerabilidades relacionadas a detenção de informações pessoais (CALO, 2017; ZARSKY, 2016), e, neste contexto, o princípio poderia ser visto como uma tentativa de propiciar um escudo a titulares dados que se relaciona a *affordances* em torno do direito à privacidade, ao esconder informações pessoais que poderiam representar vulnerabilidades (CALO, 2017). Adicionalmente, sob o ponto de vista da segurança cibernética, também é factível afirmar que, quanto menos informações pessoais um controlador hipotético possua, menos informações estão vulneráveis a quaisquer tipos de ataques ou outras formas de incidentes que os tornem disponíveis para terceiros que tenham a intenção e a possibilidade de vulnerabilizar ou explorar vulnerabilidades de titulares de dados (ZARSKY, 2016). Com relação a argumentos que poderiam embasar o princípio sob um ponto de vista teórico, Zarsky (2016) observa, citando Solove (2001), que há fundamentos razoáveis para supor que a mera detenção de informações acerca de um indivíduo por um agente de tratamento de dados teria o potencial de gerar, no titular destes dados, maior ansiedade e minar em algum grau a sua autonomia. O princípio da minimização endereça estas questões, que, por outro lado, podem ser endereçadas por meio de medidas *ex post* voltadas a proteger titulares de dados de usos abusivos e proteções insuficientes aos dados por parte de agentes de tratamento de dados, que seriam punidos após o fato por ações ou omissões que levassem à vulnerabilização de titulares (ZARSKY, 2016), como, por exemplo, o emprego de medidas de segurança aquém do desejável.

Os conflitos entre o princípio da minimização e as práticas de análise de *big data* são notáveis e recorrentemente objeto de comentários e estudos (ZARSKY, 2016). De fato, é marcante no cenário das iniciativas que dependem deste tipo de análise uma busca pela

obtenção de *mais* dados, e de mantê-los pelo período em que for possível, dados os custos de análise e coleta, com vistas a, a partir deles, gerar conhecimento valoroso por meio do emprego de técnicas existentes, ou que venham a ser desenvolvidas pelo próprio agente de tratamento ou por terceiros. Zarsky (2016, p. 1011) anota que os

> aprimoramentos constantes na ciência de dados e áreas correlatas podem gerar a crença de que o amanhã guarda uma grande promessa quanto ao que podemos encontrar enquanto analisamos dados existentes. [...] Em tese, ao menos, com mais dados deverá haver maior conhecimento, e com isso maiores benefícios para empresas e potencialmente para a sociedade em geral [...] Diligentemente aplicar o princípio da minimização limitará o sucesso de iniciativas de 'big data', ao passo que minará a sua utilidade, com justificativas possivelmente apenas limitadas para tanto. (ZARSKY, 2016, p. 1011)

Zarsky (2016) argumenta que, diante destes fatos, as imposições legais relativas à minimização dos dados devem ser reconsideradas. Possivelmente seria mais adequado uma regulação que *ex post* se voltasse a deter e punir práticas inaceitáveis com dados pessoais, enquanto viabilizaria as ricas análises de dados ora discutidas.

4.1.3 Categorias especiais

O GDPR criou uma categoria especial de dados pessoais, que se chamou "dados pessoais sensíveis", aos quais é atribuído maior grau de proteção jurídica por se compreender que representam maior grau de possibilidades de criação e exploração de vulnerabilidades dos titulares a que se referem. Estes são aqueles dados que revelem a origem racial ou étnica de uma pessoa, suas opiniões políticas, convicções religiosas ou filosóficas, participação em sindicatos, dados relacionados à saúde ou à vida sexual, dados genéticos, dados biométricos com a finalidade de identificar unicamente uma pessoa, e dados referentes à orientação sexual, conforme Article 9 do GDPR. A LGPD adota um regime similar, e no inciso II do seu artigo 5° define dado pessoal sensível como aquele "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" (BRASIL, 2018).

Atividades de tratamento de dados envolvendo este tipo de informação são possíveis, mas sujeitas a imposições legais mais restritivas, especialmente com relação às bases legais autorizadoras que podem ser empregadas para assegurar a licitude da atividade. No caso de dados pessoais sensíveis, o consentimento adquire centralidade e recebe mais adjetivações, i.e.,

contém requisitos mais estritos para a sua validade, como a necessidade de que a informação a seu respeito seja fornecida ao titular de forma destacada. Ademais, não são admitidas atividades de tratamento de dados com base no legítimo interesse do controlador em nenhuma hipótese quando se trata de dados pessoais sensíveis, tampouco com fundamento no cumprimento de um contrato. Zarsky (2016, p. 1012) observa que a justificativa para a declinação deste maior grau de proteção jurídica é intuitiva: estas categorias constituem aqueles tipos de informações que os indivíduos geralmente consideram mais privadas, e a sua ampla divulgação, independentemente do contexto, apresenta o potencial de causar ansiedade, discriminação e danos concretos. Talvez paradoxalmente, esta maior proteção se reflete, especialmente no caso da LGPD, em uma maior preponderância para o papel do consentimento do titular na validação da atividade de tratamento, o que, diante dos problemas observado por estudos como o de OOIJEN e VRABEC (2019), deve representar um sinal de alerta.

No entanto, no contexto do *big data*, a pertinência do erguimento de categorias especiais de dados pessoais, sob uma ótica instrumental, é duvidosa. Especificamente, Zarsky (2016, p. 1013) observa que uma atividade de análise de dados que se utilize apenas de dados pessoais "comuns" pode facilmente, ao final, se enquadrar em uma das categorias especiais. Um exemplo elementar: dados de compras podem permitir a dedução de dados relacionados à saúde. Ademais, o fato de que conjuntos de dados são frequentemente combinados, originando fenômeno que Solove (2013) chamou de *aggregation effect*, representa inúmeras possibilidades de que dados "comuns" originem dados de categorias especiais a partir de práticas normais, cotidianas e úteis de análises de dados, inclusive de forma inadvertida, levando a um inchaço das categorias especiais de dados.

Por um lado, o fato de que dados "comuns", no âmbito de práticas corriqueiras com dados, tornam-se frequentemente dados de categoria especial traz preocupações, custos e inseguranças para todos os que dependem de atividades de análise de dados, que não podem determinar com precisão o regime jurídico aplicável à atividade que desejam realizar. Este problema é maior para agentes de tratamento de dados menores e com menor capacidade econômica. Por outro, as características do *big data* aparentemente obscurecem ou eliminam as distinções entre as categorias; afinal, se praticamente todo tipo de dado pode produzir um dado especial, não haveria razão para a distinção, que se afigura, neste contexto, artificial. Com isso, o regime jurídico do GDPR não apenas minaria possibilidades no âmbito da análise de dados, como a disponibilidade deste tipo de análise está exercendo o efeito de minar a própria distinção legal relativamente a categorias especiais de dados especiais (ZARSKY, 2016, p. 1013). Zarsky

(2016, p. 1013) observa que a proteção conferida a categorias especiais exerceria, com isso, um papel preponderantemente simbólico, ao sinalizar que atividades com este tipo de dado teria o potencial de gerar maiores danos, e assim deveriam ser levadas a cabo diante de maiores cuidados. No entanto, a realidade das práticas relacionadas a *big data* pode também demonstrar a impropriedade da distinção, mesmo em um nível simbólico: conforme Rouvroy (2016 *apud* ZARSKY, 2016, p. 1014), hodiernamente as práticas de discriminação ocorrem de forma distinta, e não dependem necessariamente de intento discriminatório. De fato, podemos deduzir do trabalho de Calo (2017) que as práticas atuais relacionadas à criação e exploração de vulnerabilidades a partir do uso de informação não se relacionam, necessariamente, com os aspectos delineados pelas categorias especiais. Schwartz e Solove (2011) também observam que, na contemporaneidade, atividades com dados que causam danos a titulares não estão, no mais das vezes, relacionados aspectos usualmente tidos por discriminatórios. Neste ponto, podemos nos recordar da conceituação de problema que decorre da metáfora kafkiana, conforme apresentada por Solove (2006): Josef K. não foi preso e executado em decorrência de convicções políticas ou filosóficas, mas por um tratamento desumanizado e utilização de informações a seu respeito sem qualquer procedimento adequado. Nestas circunstâncias, a própria utilidade da distinção, sob o ponto de vista da proteção dos titulares das informações, é duvidosa.

Em suma, os fenômenos relacionados à análise de *big data* substancialmente denotam que a lógica e a utilidade em se valer de categorias especiais no regime jurídico de proteção de dados pessoais não se sustentam. Sob um ponto de vista prático, a distinção traz custos regulatórios altos e desnecessários. Além disso, gera um grau elevado de incerteza. Todos estes aspectos afetam em maior grau pequenas empresas relativamente àquelas já estabelecidas no mercado, o que representa um fator que favorece a concentração de mercado no setor, o que, por sua vez, se afigura como um dos fatores mais relevantes na vulnerabilização de titulares de dados. Além disso, considerando a justificativa simbólica de se inserir a distinção como forma de denotar maior grau de vulnerabilidade a que estão sujeitos os titulares de dados reputados sensíveis, "é importante enfatizar que outras razões simbólicas para abandonar o uso de categorias especiais. Se quase todos os dados podem cair na categoria de dados 'especiais', o sinal e a mensagem que este *framework* regulatório transmitem com relação ao nível de privacidade devido em razão da especialidade é subsequentemente diluído" (ZARSKY, 2016, p. 1014). De fato, esta afirmação é consistente com estudos acerca do papel de sinais no processo interpretativo, e nos permite estender o argumento para afirmar que, sob a ótica dos

titulares de dados, a existência destas categorias especiais podem funcionar como um sinal de controle (INNES, 2004) que falsamente transmite aos titulares sensação de segurança, seja pela crença de que dados sensíveis serão tratados com maior zelo, seja pela equivocada percepção de que a cessão de dados "comuns" não representaria os mesmos riscos do que ter dados sensíveis sob tratamento – o que, conforme argumentamos acima, não é verdadeiro, seja porque outros tipos de informação são igualmente capazes de gerar discriminação, seja porque dados "comuns" podem ultimamente levar à revelação de dados de categorias especiais.

4.1.4 Decisões automatizadas

O Artigo 22 do GDPR chama atenção por sua regra reputada bastante singular com relação a processos de tomada de decisões por meio de sistemas totalmente automatizados e que substancialmente impactem indivíduos, como em análises de crédito ou seleções de currículos de interessados em vagas de emprego (ZARSKY, 2016). Zarsky (2016) nota que a norma, que não encontra par, por exemplo, no ordenamento jurídico norte-americano, assegura aos europeus o direito de *não serem sujeitos* a tais processos, senão nos casos excepcionados pelo Regulamento, como nos casos em que há consentimento expresso do titular para tanto; quando necessário para a realização de um contrato; e outras que podem ser dispostas pelos Estados Membros. Mesmo nos casos excepcionados, ao titular são assegurados importantes direitos com relação a decisões automatizadas, como o de "obter intervenção humana" e o de "contestar a decisão" expressando o seu ponto de vista, além de terem direito de acessar os dados pessoais que embasaram o processo para fins de realizar esta contestação (Zarsky, 2016, p. 1016). Além disso, o Regulamento prevê que o titular terá direito a ser informado em qualquer hipótese em que este tipo de atividade de tratamento de dados pessoais ocorra com relação a si, e ainda de receber "informações significativas relativamente à lógica envolvida, bem como o significado e as consequências antevistas desta atividade de tratamento para o titular de dados" (ZARSKY, 2016, p. 1016). No caso da LGPD, esta regra também existe e está disposta no artigo 20 na forma de um "direito a solicitar a revisão de decisões tomadas unicamente com base em tratamento automatizado de dados pessoais que afetem seus interesses" (BRASIL, 2018). O parágrafo primeiro do referido artigo dispõe que, observados os segredos comercial e industrial, o "controlador deverá fornecer, sempre que solicitadas, informações claras e adequadas a respeito dos critérios e dos procedimentos utilizados para a decisão automatizada" (BRASIL, 2018).

Zarsky (2016) cita diversas justificativas possíveis para a inserção desta regra e similares. Primeiramente, poder-se-ia "conectar este direito a uma noção de honra e respeito; quando defrontado com decisões relevantes, um ser humano deve ser tratado com a dignidade de ter um tomador de decisão humano endereçando a sua questão pessoal" (ZARSKY, 2016, p. 1017). No vocabulário jurídico brasileiro, poderíamos associar este conceito ao princípio da dignidade da pessoa humana. Em segundo lugar, "o fato de que estes processos automatizados ocorrem sem prover suficientes *insights* àqueles por eles afetados mina um direito ao 'devido processo'", e levanta preocupações relacionadas à ocorrência de erros e discriminações (ZARSKY, 2016, p. 1017). Sob outra perspectiva, esta regra pode refletir a "desconfiança de humanos para com sistemas computadorizados e máquinas de forma ampla" (ZARSKY, 2016, p. 1017), nos contornos delineados por Sunstein (2017).

Fora argumentado que a norma, em conjunto com os "direitos de acesso" que confere a titulares de dados, impacta diretamente práticas de *big data*, e que o atingimento de seus objetivos no contexto do *big data* seria irrealista e paradoxal. A tensão entre as práticas de *big data* e a regra relativa a decisões automatizadas pode ser observada em diferentes aspectos. Primeiramente, impedir decisões automatizadas impede diversas práticas de *big data* relevantes e que se fiam em sistemas automatizados para a tomada de decisões, inclusive com importantes ganhos em termos de eficiência e acurácia. Em segundo lugar, os direitos de acesso concedidos a titulares para fins de questionamento e conhecimento destas decisões frequentemente representam obrigações de difícil ou de impossível cumprimento por controladores que se utilizem destas práticas: para cumpri-las, os processos de *big data* precisam ser levados a cabo de forma a ser possível a sua interpretação, i.e., explicação para o indivíduo solicitante de informações a respeito, e a busca constante por este tipo de "interpretabilidade" possivelmente levará os responsáveis por desenvolver tais sistemas automatizados a sacrificar a eficiência ou a precisão do sistema, o que por sua vez poderá representar processos pouco eficientes e que deixam de entregar à sociedade aquilo que teriam o potencial de fazê-lo, sem necessariamente incrementar as possibilidades de controle individual.

Todos estes problemas, conforme se observa, decorrem de uma compreensão ou tratamento equivocado do problema que se endereça. Uma vez que a coleta, geração e utilização de informações em meios digitais não ocorre necessariamente da mesma forma com que ocorre em suporte físicos, e pelo fato de que práticas que envolvem a geração de informações fazerem parte do quotidiano e da estrutura da sociedade contemporânea, o regulamento em torno de proteção de dados pessoais que se coloca como antagonista de práticas de tratamento de dados

e afasta o titular destas práticas não é capaz de exercer o papel do direito em tutelar liberdades e garantias fundamentais, ao passo que resguarda o espaço de atuação dos agentes com maior poder neste âmbito. A delimitação do âmbito de aplicação da norma, e a compreensão da materialidade que visa regulamentar, apresentam-se como aspectos centrais; por outro lado, conforme indicam pesquisas recentes, representam dois aspectos de um mesmo problema.

4.2 Mercados de dados pessoais e regulação jurídica

A análise de Zarsky (2016), sobretudo normativa, pode ser enriquecida a partir de perspectivas que levam em consideração os impactos da regulação sobre proteção de dados pessoais sobre os agentes econômicos, e, sobretudo, sobre as atividades que estes desempenham sobre dados pessoais. A abertura do direito para as ciências e para a multidisciplinariedade, como forma de aprimorar as soluções regulatórias e jurídicas, deve ocorrer não apenas sob o ponto de vista da incorporação, no processo decisório, de estudos científicos sobre o objeto de atuação; é importante que o Direito esteja aberto para incorporar visões dos processos jurídicos por parte de outros campos de conhecimento (CARDOSO et al., 2018).

A partir de perspectiva de análise econômica do Direito, Gal e Aviv (2020) avaliaram os efeitos do GDPR sobre agentes econômicos que dependem de atividades com dados pessoais. O mercado relacionado a dados consiste em diversos elos ao longo da cadeia de geração de valor, como coleta, processamento, e uso de informação e conhecimento gerado a partir de dados, armazenamento, compartilhamento e destruição. Destes, considera-se principais os três primeiros: a coleta de dados se refere à extração de dados e à sua datificação, que representa atividades como a gravação em suporte digital, agregação e organização da informação em formato que viabilize o uso para data mining, inclusive a sua transferência a servidores conectados à rede mundial de computadores. O processamento relaciona-se à otimização, remoção de ruídos, *parsing*, ou combinação de diferentes conjuntos de dados, com a finalidade de organizar os dados para futuras extrações e buscas por correlações. Esta etapa, portanto, tem o potencial de transformar dados crus em informação, definida como dados em contexto, e pode incluir também a análise de dados, desta forma criando conhecimento. Por fim, o uso é definido como o emprego de informações ou conhecimento baseados em dados para atividades preditivas e de tomada de decisão em contextos específicos. Além de poder levar a processos, produtos, serviços e predições melhores ou inovadores, este elo da cadeia tem uma dimensão

recíproca interna: através do *machine learning*, informações relacionadas à performance pretérita de predições podem ser utilizadas para "ensinar" o algoritmo para realizar melhores predições no futuro, levando ao que Surden (2014, pp. 89-95) denominou neste contexto de feedback loop (GAL e AVIV, 2020, pp. 8-9).

Gal e Aviv (2020) basearam-se, inicialmente, em três pressupostos acerca das características de mercados relacionados a dados, que em grande medida podem ser associados às características da informação em meio digital identificadas por Zarsky (2016). O primeiro pressuposto se relaciona à importância dos dados pessoais. Como enfatizado por diversos estudos, com os dados se transformando rapidamente no sangue vital da economia digital, as formas e a eficiência de seus usos afetam significativamente o bem-estar, tanto privado, quanto social. É notório atualmente o papel relevante de predições baseadas em padrões e correlações identificadas a partir de dados, inclusive em aspectos que representam preocupações constantes e elementares da sociedade, como educação, saúde, transporte e sustentabilidade. Em algumas searas, a coleta e o uso de dados pessoais são intrínsecos ou necessários para a atividade. O segundo pressuposto é o de que, a despeito das diferentes necessidades de diferentes atores econômicos com relação a dados pessoais, para muitos de seus usos a qualidade da informação e do conhecimento que podem ser extraídos estão correlacionados estreitamente com quatro elementos dos dados: seu volume (a quantidade de dados utilizados na análise), sua variedade (a diversidade de suas fontes), sua veracidade (acurácia e confiabilidade), e sua velocidade (relativo a ser atualizado para os fins a que se propõe, chamado pela doutrina em língua inglesa de *freshness*). Para embasar esta correlação, há três razões cumulativas.

A primeira razão é comumente caracterizada a partir da noção de economias de escala, o que implica no fato de que são necessários, em princípio, grande escala e alcance da atividade relacionada a dados pessoais para que esta seja significativa e possa se inserir na sociedade, através do mercado ou política pública, de forma competitiva e atrativa. Esta necessidade decorre diretamente da importância dos quatro elementos referidos para a qualidade do conhecimento alcançado a partir de dados. A segunda razão se relaciona ao que se chama de feedback loop quanto à atividade de algoritmos utilizados para análise de dados, que se beneficiam diretamente do aprendizado obtido a partir de predições pretéritas; assim, algoritmos com acesso a melhores bases dados (conforme as quatro qualidades mencionadas) tornam-se progressivamente melhores, pois capazes de realizar melhores predições. Vale dizer que esta característica também é responsável por muitas das preocupações a respeito do intrínseco potencial discriminatório pelo emprego de técnicas de aprendizado por máquina a

partir de dados históricos (EDWARDS e VEALE, 2017, p. 31). A terceira é a constatação de que as qualidades de um conjunto de dados podem criar externalidades positivas também com relação a outros conjuntos de dados, o que decorre do efeito denominado de "aprendizado por transferência", ou seja, o fato de que um algoritmo pode "aprender" a partir de seu emprego em um conjunto de dados de alto valor para realizar determinada tarefa, que após poderá ser realizada sobre outros conjuntos de dados.

O terceiro pressuposto é o de que em muitas configurações de mercados relacionados a dados, o papel do compartilhamento de dados é central na realização dos benefícios potenciais baseados em dados. A razão para tanto é o fato de que muitos dados são coletados em um sistema que é amplamente modular e distribuído (GAL e AVIV, 2020, p. 10). Representativamente, em 2020, 30 bilhões de dispositivos de Internet das coisas, controlados por atores diversos, estão conectados à internet, constantemente coletando e utilizando dados. Incentivos ao compartilhamento de dados podem ampliar a competitividade em um mercado notoriamente marcado pela concentração. Por outro lado, o compartilhamento de dados pessoais também pode ter um efeito negativo na sociedade, como discriminação de preços, danos morais, danos à esfera de autonomia e à liberdade de expressão, e, assim, criam trade-offs complexos e ambíguos que merecem cuidadoso e consciente balanceamento diante de interesses relevantes em conflito.

Pode-se dizer que dados, após coletados, são bens não-rivais (LOI et al., 2020). Por outro lado, em muitos mercados, barreiras de entrada consideráveis existem com relação à coleta de dados (RUBINFELD e GAL, 2017). Neste cenário, o seu compartilhamento poderia empoderar mais pessoas, físicas e jurídicas, para usar dados e criar sinergias entre conjuntos de dados, diante das quatro características elementares de conjuntos de dados valorosos (GAL e AVIV, 2020). O compartilhamento de dados pessoais tem o potencial de estender os benefícios de uma atividade de *escala* com dados para atores sociais e econômicos que, sozinhos, não têm a capacidade de obter e processar dados de forma a gerar conhecimento valoroso. Além disso, no contexto de investigação de cunho exploratório acerca do impacto da inteligência artificial sobre a inovação, foi argumentado que barreiras ao compartilhamento de dados podem resultar em um processo de balcanização de dados em determinados setores econômicos ou empresas, o que, além de impedir a inovação, reduz incentivos para a melhora de ferramentas analíticas (COCKBURN, 2019). Esse potencial processo de balcanização pode ser relacionado, ainda, à indiferença marcante em muitos contextos de tomada de decisões embasadas em dados, que resulta em amplificação da situação de vulnerabilidade de pessoas dependentes em tais

processos (SOLOVE, 2006). Por outro lado, o foco na utilização dos dados também fornece as ferramentas para uma melhor compreensão acerca de suas externalidades, positivas e negativas, a partir de um olhar que vai além da atividade que originou a coleta, considerando que, sob certo sentido, estas podem ser consideradas atividades secundárias nos mercados relacionados a dados pessoais.

Uma das importantes desigualdades quanto às relações que envolvem atividades de tratamento de dados e as variadas potencialidades de emprego do conhecimento deles decorrente se refere à capacidade de coletar e controlar estes ativos, i.e., informação, que é restrita a poucos *gatekeepers* da internet (Google, Facebook etc.) que estão posicionados unicamente para realizar esta função. Esta posição singular do *gatekeeper* decorre de uma combinação de vários *efeitos de rede* que tornam difícil, senão impossível, competir contra a primeira companhia que deles se beneficia em um dado contexto. Por exemplo, o serviço de busca do Google se beneficia de diversos destes efeitos, que atuam de forma interrelacionada:

> Marketplace network effects (advertisers affiliated with Google can access individuals with the best profiles, while each advertiser contributes to profiling), data network effects (more data makes it easier to build an ecosystem of services around each user), recruiting network effects (the behaviour of users tells Google which search results are selected after typing a given search key, thus helping Google identify the most fitting search results). Due to these network effects, internet services markets tend to be winner-takes-all. (LOI et al., 2020, p. 3)

Os usuários destes serviços, tipicamente, não são capazes de conceber os dados que produzem, as formas com que são combinados e ulteriormente utilizados, e o valor que geram e que têm o potencial de gerar, sendo-lhes oferecidas apenas "intangíveis e inquantificáveis transações com seus dados, o que sistemática e significantemente dificulta o surgimento de arranjos econômicos mais favoráveis" (HAYNES e CAROLYN NGUYEN, 2014, p. 8).

Os riscos relativos a controlar dados pessoais, por seu turno, são significativos e se revelam cada vez mais elevados. De fato, se dados são ativos, também são, sob outro aspecto, passivos, no sentido de que geram custos de diversas naturezas e que podem ser categorizados de diferentes formas, e.g., efetivos ou contingentes. Além disso, devemos considerar que inovações tecnológicas, em geral, representam, sob o ponto de vista econômico, bens públicos, o que representa uma falta de incentivos naturais para o engajamento no seu desenvolvimento (FISHER, 2001). Diante deste fato, qualquer incremento nos custos de se dedicar a inovações com relação a práticas de análises de dados deverá ser sopesada com aos benefícios potenciais da limitação. A intervenção jurídica tem se mostrado significativa quanto aos incentivos que

representa para agentes envolvidos em atividades de tratamento de dados, ou agentes interessados em realizar atividades desta natureza, ou que poderiam se beneficiar de atividade com conjuntos de dados. A questão principal que se coloca é que tipo de efeito esta intervenção tem causado, e ainda se ela afeta de maneiras similares ou diversas os diferentes atores envolvidos nestas atividades.

Zuboff (2019, p. 12) argumenta que a lógica de acumulação de dados pessoais a partir do excesso preditivo-comportamental dos dados gerados permite a indução de imperativos econômicos chave. Prazeres (2021), em argumento que denota a *antifragilidade* das atividades econômicas desempenhadas por grandes empresas de tecnologia – que antes *determinam* os rumos das jurisdições nacionais, do que são por eles determinados[8] – , observa que

> não apenas se vislumbra, internamente ao sistema econômico, a ascensão de uma nova lógica de acumulação, em que o domínio e acesso às estruturas cibernéticas – *e, por consequência, aos dados nelas armazenados* – importa, como se enfrenta uma possível crise de expectativas do direito [...]. A falha no contingenciamento do movimento expansionista do mercado se torna bem flagrante quando se atenta para o *modus operandi* das grandes empresas de tecnologia, consistente em fazer incursão em âmbitos privados e não protegidos até que alguma resistência seja encontrada. (PRAZERES, 2021, p. 18)

Sob lentes mais pragmáticas, Gal e Aviv (2020) reuniram as características essenciais dos atuais mercados de dados pessoais, com a finalidade de avaliar os eventuais efeitos do GDPR neste âmbito. Quaisquer pessoas, físicas ou jurídicas, que necessitem de dados pessoais como input para as suas operações precisam de uma estratégia para os obter e processar. A partir de entrevistas com participantes do mercado, GAL e AVIV (2020, p. 11) identificaram as cinco principais estratégias empregadas por estes participantes para obter dados pessoais relevantes para suas atividades, a saber:

1. Coletar organicamente dados pessoais (“first-party data”, ou “dados de primeira mão”), ou seja, o dado é coletado pelo próprio agente de tratamento de dados diretamente do titular dos dados pessoais;

[8] Exemplos não faltam: os casos do Uber, Airbnb e iFood, por exemplo, demonstram como estes agentes têm sido capazes de (1) determinar expectativas da sociedade, e (2) moldar as legislações nacionais para acomodá-los. Vale dizer que esta acomodação, frequentemente, *fecha* o espaço para outras iniciativas similares. Em outras palavras, reforça a lógica do *winner takes it all* deste mercado, e artificialmente cria monopólios ou oligopólios, a seu turno enfraquecendo a capacidade de o direito tutelar direitos fundamentais, o que é agravado pelo forte caráter normativo dos sistemas em que tais empresas atuam.

2. Unir-se a uma entidade que possui os dados necessários, e os utilizar em suas próprias atividades;

3. Comprar ou receber os dados de um fornecedor externo (“third-party data”, ou “dados de segunda mão”);

4. Tornar-se parte de uma joint venture através da qual empresas realizem um pool de seus dados, ou de seu conhecimento baseado em dados, para propósitos específicos pré-determinados;

5. Comprar ou receber conhecimento baseado em dados (ao invés dos dados em si), ou dados agregados, de um fornecedor externo (“knowledge broker”, ou “corretor de conhecimento”). Destaca-se que esta modalidade não é sujeita ao GDPR.

Com relação aos chamados “dados de segunda mão”, destaca-se a ocorrência deste tipo de compartilhamento por meio das denominadas *application programming interfaces* (APIs), que são interfaces ou protocolos de comunicação entre um cliente e um servidor, por meio do qual o servidor inicia uma ação previamente definida, inclusive, neste caso, prover dados, em resposta a uma requisição enviada pelo cliente solicitando dados em um formato específico (GAL e AVIV, 2020, p. 11). Este mecanismo é utilizado, por exemplo, para que empresas acessem bancos de dados de empresas como Facebook e Google para a realização de propósitos específicos, notoriamente autenticação ou obtenção de dados pessoais para fins relacionados ao serviço por ela fornecido, a partir da autorização do titular de dados para que seja feita a requisição. Isso ocorre, por exemplo, quando nos cadastramos em serviços online por meio de funcionalidades como “utilizar o perfil do Facebook” para extração das informações cadastrais, ao invés de realizar um cadastro autônomo junto à empresa fornecedora do serviço desejado. Por outro lado, o Twitter é um exemplo de plataforma digital que fornece uma API para que se obtenha, em formato legível por máquina e de forma massiva, tweets publicados na plataforma, para que o interessado possa, por exemplo, realizar análises relacionadas à frequência da ocorrência de determinado termo ou expressão em tweets publicados a partir de critérios determinados (localização geográfica, períodos de tempo, grupos de usuários etc.).

As opções referidas nos itens 1 e 5 geralmente não permitem a realização de sinergias de dados entre empresas (GAL e AVIV, 2020, p. 12). Destaca-se ainda que o item 5 se refere à atividade não abrangida pelo GDPR ou LGPD, e, ademais, tem a característica de, ao contrário das demais opções referidas para a obtenção de dados por empresas, não permitir o uso dos

dados internamente ou o seu reaproveitamento, pois estes não lhe são entregues, mas somente o conhecimento neles baseado. Exemplo deste mercado é um dos serviços oferecidos pelo Google: o usuário acessa o serviço e insere um query, ou seja, um termo de busca, que é enviado por meio de uma requisição ao servidor, que responde com base no banco de dados do Google, porém jamais expondo os dados que serviram de base para as respostas (GAL e AVIV, 2020, p. 11).

A escolha entre estas opções é baseada em critérios relacionados ao seu custo-benefício relativo, e em sua viabilidade e escalabilidade, que, por sua vez, são afetados por uma combinação de barreiras tecnológicas, financeiras, estratégicas e legais (GAL e AVIV, 2020, p. 12). Barreiras tecnológicas são os fatores que impedem a coleta ou o compartilhamento de dados sob um ponto de vista técnico, como barreiras à interoperabilidade entre bases de dados que são organizadas por entes diversos, de acordo com lógicas diversas. Algumas barreiras tecnológicas podem ser superadas, porém sob determinado custo; por outro lado, outras são proibitivas. Barreiras financeiras são fatores que impedem a obtenção de dados com custo-benefício positivo. Barreiras estratégicas estão associadas àquelas empregadas pelos detentores de dados para manter o seu poder de mercado. As barreiras legais, por fim, são impostas pela legislação com relação a atividades relacionadas com dados pessoais, notadamente a sua coleta, processamento e utilização (GAL e AVIV, 2020, pp. 12-3).

Este cenário é consistente com o que foi desenvolvido por Lessig (1996) quanto aos modelos regulatórios da conduta humana operantes sobre os membros da sociedade, e de como a análise da interação entre o direito e outros sistemas de incentivos pode lançar luz sobre difíceis escolhas com que o campo jurídico se depara desde o advento e popularização da rede mundial de computadores. Gal e Aviv (2020), em sua análise, tratam as barreiras não-legais como dadas, dando enfoque a como as barreiras legais afetam a escolha entre os cinco modelos descritos para a obtenção de dados pessoais por empresas que deles necessitam para suas atividades. Em alguns casos, o regime jurídico de proteção de dados mostrou efeito decisivo, reforçando o notório poder regulatório estatal em moldar potenciais interações de mercado, e a necessidade de cuidadoso balanceamento ao erigir disciplinas jurídicas abrangentes.

Em seguida, destacaremos os aspectos observados por Gal e Aviv (2020) quanto às formas com que o GDPR altera aspectos das atividades que envolvem tratamento de dados pessoais.

4.3 Obrigações legais com relação a atividades com dados pessoais

Legislações em matéria de proteção de dados podem impor obrigações relativas à coleta, processamento, armazenamento, uso, compartilhamento, e outras atividades com dados pessoais, como é o caso do GDPR e da LGPD. Estas limitações legais, por sua vez, afetam as escolhas de empresas ou outras pessoas que necessitam de dados pessoais para desempenhar as suas atividades com relação aos modelos de negócios disponíveis para a obtenção destes dados.

Parte-se do pressuposto de que, quando dados pessoais forem necessários para uma determinada atividade, o interessado buscará a forma mais eficiente e menos custosa de os obter. Por exemplo, quando o compartilhamento de dados for muito custoso ou não for lícito, haverá maiores incentivos para que se colete dados internamente, e vice-versa, o que representa o *trade-off* implícito na escolha entre a obtenção de produtos e serviços interna ou externamente ao organismo responsável pela atividade (GAL e AVIV, 2020). Com relação aos regimes jurídicos de proteção de dados, quanto mais fortes as sanções (contratuais, legais, reputacionais etc.) para a falta de adequação à norma, mais importante se torna para empresas que necessitem de dados pessoais a adoção de modelos de negócios para a sua obtenção que assegurem o reconhecimento do *compliance*, pois isto afeta diretamente a relação de custo, eficiência e risco da operação.

São apresentadas algumas das formas com que regimes jurídicos afetam estas escolhas, conforme análise de GAL e AVIV (2020, p. 13). Embora o foco seja nas barreiras jurídicas, são mencionadas barreiras tecnológicas, financeiras e estratégicas que sejam relevantes ao contexto discutido. São identificados cinco casos, sem pretensão de exaustão da matéria, que se reputa com maior potencial de influenciar nas escolhas realizadas neste âmbito. Para tanto, propõe-se avaliar, nestes cinco contextos, os diferentes custos, riscos, limitações e benefícios que o regime jurídico de proteção de dados impõe com relação às possibilidades de obtenção de dados pessoais necessários para atividades legítimas, identificando, por fim, como ele afeta estas escolhas.

4.3.1 Garantia da legalidade da atividade de tratamento de dados

No caso do regime jurídico europeu, GAL e AVIV (2020, p. 14) observaram que toda atividade de tratamento de dados pessoais deve observar quatro elementos fundamentais: os

dados devem ser tratados de forma lícita, de boa-fé[9], e transparente, o que constitui no regime jurídico europeu o *lawful basis*; devem ser coletados apenas para finalidades especificadas, explícitas e legítimas (limitação das finalidades admitidas); deve ser limitada aos dados necessários para as finalidades do agente de tratamento de dados, o que se chamou de *minimização;* e devem ser mantidos atualizados e corretos.

A *licitude* da atividade, tanto no regime jurídico de proteção de dados europeu, quanto no brasileiro, é baseada na observância de uma das bases legais alternativas trazidas pela norma. No caso da LGPD, estas estão no artigo 7° do diploma, com relação aos dados pessoais em geral, e no artigo 11, relativamente às bases legais admitidas para o tratamento de dados pessoais sensíveis, estes definidos no inciso II do artigo 5° da LGPD. Tanto no contexto europeu, conforme se depreende da análise de GAL e AVIV (2020, p. 15), quanto na disciplina de proteção de dados pessoais vigente no Brasil ao tempo da escrita desta dissertação, as bases legais mais relevantes para a investigação ora empreendida são o consentimento do titular e o legítimo interesse do controlador.

Para os fins deste estudo, quanto ao consentimento, destaca-se que este, factualmente, se refere tanto à utilização dos dados pessoais, quanto à pessoa, física ou jurídica, que realizará a atividade. Assim, presume-se que um titular de dados poderia ter a inclinação de consentir para com determinado uso de um conjunto específico de seus dados pessoais, por um determinado agente, porém não apresentar a mesma inclinação com relação ao mesmo uso, dos mesmos dados, por outro agente (GAL e AVIV, 2020, p. 15). De fato, embora discussões acaloradas e importantes ocorram com relação à sua posição no sistema jurídico de proteção de dados, ao seu conteúdo normativo e outras características relevantes, pode-se dizer afirmativamente que a estrutura normativa da LGPD dispõe que o consentimento deverá ser específico e expresso. Além disso, o compartilhamento de dados pessoais é tratado pela LGPD como uma atividade de tratamento de dados *lato sensu*, o que, atuando em conjunto com o princípio da finalidade, reforça o caráter de especificidade do consentimento, neste caso sob o aspecto do agente autorizado a realizar licitamente a atividade especificada. Para que a atividade seja desempenhada, com os mesmos dados, por outro agente, isto deverá ter sido especificado quando da obtenção do consentimento sempre que este for a base legal para a atividade, o que

[9] Escrevendo em inglês, os pesquisadores se referiram a *fairness*, que, no contexto brasileiro, entendemos contemplado no artigo 6° da LGPD, que traz os princípios a serem observados por toda atividade de tratamento de dados, com papel de destaque à *boa-fé* – não hierarquia –, constante do *caput*, enquanto outros princípios constam de seus dez incisos.

coloca a aquiescência com aquele responsável pela atividade, de fato, como conteúdo da manifestação de vontade caracterizada juridicamente pelo consentimento, ao lado da finalidade da atividade, embora isso não esteja expresso no texto normativo quando trata especificamente das qualidades e características inerentes ao instituto do consentimento.

Em geral, argumenta-se que é mais fácil de obter consentimento quando este é a base legal adequada no modelo de coleta interna de dados, por duas razões principais. A primeira é que o uso dos dados pela própria pessoa que os coleta é mais direto e intuitivo, e a segunda, que o consentimento obtido por um agente para o seu uso pode ser aplicável a todas as suas divisões internas, inclusive àquelas que sejam criadas posteriormente (GAL e AVIV, 2020, p. 15). Neste sentido, em termos econômicos, não apenas a escala, mas a abrangência da atividade de determinado agente de tratamento de dados representa enorme vantagem competitiva potencial, pois ao combinar os requerimentos de consentimento para todos os seus usos podem ser reduzidos drasticamente os custos relacionados à obtenção do consentimento com as qualidades especificadas pelas normas de proteção de dados pessoais, ou seja, os custos relacionados a assegurar a licitude da atividade de tratamento. Ilustrativamente, pesquisas de mercado têm provido evidência empírica de que a natureza dos custos transacionais relativos à obtenção do consentimento no regime jurídico erigido pelo GDPR, para a realização de atividades de *marketing* digital, reduziu os entrantes no mercado e desproporcionalmente afetou a fatia de mercado daquelas empresas de menor porte que já competiam no mercado, com relação aos maiores 50 atores econômicos no setor (GAL e AVIV, 2020, p. 15).

Como exemplo da vantagem competitiva mencionada para empresas com escala e variedade de suas atividades, que necessitam de obter consentimento para assegurar a licitude da operação, tomemos o caso do Google. A empresa, por meio de sua interface, solicita ao usuário que clique no botão "*sign up*", com a notificação de que esta ação representa *consentimento* às políticas de privacidade, de dados pessoais e de *cookies* da empresa com relação a todos os dados do usuário. Os usuários podem, a qualquer momento, retirar o consentimento (*opt out*) com relação a alguns ou a todos os usos de seus dados. Observe-se que, a despeito da dicotomia bastante presente quando das discussões sobre proteção de dados quanto a sistemas de *opt in* ou *opt out,* e da adoção, pelo GDPR e pela LGPD, do primeiro[10],

[10] Ou seja, o usuário precisa consentir para uma atividade de tratamento com seus dados pessoais, não se admitindo a presunção de consentimento senão até a sua retirada para fins de aferição da licitude da operação com dados pessoais.

esta granularização[11] do consentimento requer ações por parte do usuário, enquanto *por design* resta assegurado a empresas com escala e variedade de atividades situação de maior eficiência, sob o ponto de vista estrito de seus interesses. Por outro lado, interessados em tratar dados de forma mais restrita são desproporcionalmente afetados quanto aos custos transacionais de obtenção do consentimento do titular para assegurar a licitude da atividade sob a atual disciplina de proteção de dados pessoais brasileira, como naquela vigente no continente europeu, relativamente às grandes companhias.

Tem sido observada a propensão maior dos usuários de fornecerem dados pessoais para empresas mais conhecidas no mercado ou com as quais já mantenham relação anterior, o que representa uma maior facilidade de que estas companhias obtenham consentimento dos titulares de dados para um conjunto maior de finalidades. De fato, estudo demonstrou que consumidores são mais propensos a conceder consentimento para grandes companhias, com amplo escopo de mercado, do que a empresas menos estabelecidas no mercado (GAL e AVIV, 2020, p. 16). Ademais, empresas que fornecem múltiplos serviços ou produtos têm mais "portas de entrada" por meio das quais podem obter o consentimento dos titulares para operações com dados pessoais.

Em termos exemplificativos: uma hipotética empresa G fornece múltiplos serviços (C, D e E), e por meio deles coleta dados pessoais (h, i e j) dos usuários, para finalidades diversas (Fw, Fx, Fy, Fz). Consideremos que a coleta de parte destes dados (h e i) seja intercambiável quanto a todos os serviços, i.e., podem ser coletados quando o usuário utiliza quaisquer dos serviços C, D ou E, apresentando para G o mesmo valor independentemente do serviço de origem. Os dados h e i servem às finalidades Fy e Fz. As finalidades Fy e Fz, além de relacionadas aos serviços prestados (C, D e E), são voltadas aos serviços H e I, também prestados pela empresa G, porém para público-alvo diverso – clientes que adquirem conhecimento de G, construído a partir dos dados h e i. Quando o usuário utiliza quaisquer dos serviços (C, D, E), aquiescerá com o fornecimento dos dados h e i, para as finalidades Fy e Fz, além dos dados e finalidades inerentes ao serviço que no momento utiliza.

Isso pode ser mais bem compreendido, também, sob o ponto de vista da imposição estratégica de uma limitação potencial à cognição do titular dos dados quanto à natureza da transação com que aquiesce ao franquear acesso a seus dados pessoais para uma finalidade

[11] Ou seja, um consentimento mais específico e fragmentado, e que possibilita maior personalização das escolhas do titular.

específica. Em outras palavras: a doutrina observa que a decisão, pelo titular, relacionada ao franqueamento de acesso a determinada parcela de seus dados pessoais, para uma ou mais finalidades especificadas, é contextual, e é essencialmente uma decisão acerca do risco envolvido na operação (Solove, 2020). Este risco, primeiramente, deve ser percebido pelo indivíduo, e adentrar a sua esfera de consciência. Será, então, sopesado com outros riscos e benefícios percebidos da operação. O sujeito deverá, ainda, realizar juízo acerca da confiabilidade das predições de risco. Esta não é matéria trivial, como veremos em capítulo adiante. O que se destaca, neste momento, é que as possibilidades de coleta de dados valiosos são tão maiores quanto maiores sejam o tamanho, a variedade e a abrangência das atividades de um *player*, o que lhe assegura progressivas vantagens, favorecendo a concentração de mercado, de dados, e de poder sempre que as condições não favoreçam a *desconcentração*.

O interesse legítimo é base legal alternativa ao consentimento capaz se assegurar a licitude da atividade de tratamento. O GDPR, como a LGPD, determinam que, nestes casos, o controlador interessado em embasar sua operação com dados pessoais no seu legítimo interesse deverá realizar uma avaliação tripartite para assegurar a conformidade do alegado interesse com os direitos dos usuários. Primeiramente, um interesse que possa ser reputado legítimo deve ser identificado. Após, deve ser demonstrado que o tratamento de dados pessoais é necessário para os objetivos representados pelo interesse. Por fim, os direitos individuais dos titulares devem ser balanceados com relação ao legítimo interesse do controlador, através da estrutura da máxima da proporcionalidade. Se, por um lado, o interesse legítimo constitui uma das bases legais mais flexíveis viabilizadas pelo GDPR e pela LGPD, os limites dessa base legal para assegurar licitude da operação com dados pessoais ainda é imprecisa, especialmente no Brasil, onde houve menos oportunidades de ser interpretada em ocorrências de conflitos de interesses. A título ilustrativo da incerteza relativa a esta base legal no continente europeu, onde o debate já possui alguma maturidade, tomemos o caso da agência *online* de viagens *TravelBird*, que, quando de seu processo de falência, vendeu ao competidor *Secret Escapes* a sua base de dados de clientes – ou seja, dados pessoais de seus clientes. O comprador informou previamente a todos os titulares, concedendo-lhes duas semanas para se opor ao potencial novo agente de tratamento a cargo da atividade, ressalvando que a finalidade do uso dos dados seria a mesma para os quais os titulares haviam consentido, e conforme as políticas de privacidade com as quais haviam aquiescido, ou seja, aquela finalidade e aquela política de privacidade da empresa que vendeu a base de dados. As discussões envolvem aspectos como a base legal autorizadora da venda (consentimento ou legítimo interesse), e a relação de eventuais regras jurídicas

relativas à obrigação de empresa solvente adequadamente gerir seus ativos na ocorrência da liquidação. Inicialmente, não houve oposição do *European Data Protection Board* quanto à operação da TravelBird (GAL e AVIV, 2020, p. 18). Vê-se que não apenas os custos com a formação ou utilização de uma base de dados podem influir nas decisões de atores econômicos, mas o próprio caráter de ativo ou passivo de uma coleção de dados pessoais gera impactos na atividade que não são ignorados por administradores e por normas inicialmente descorrelacionadas com a disciplina de proteção de dados pessoais.

4.3.2 Assegurar o *compliance* por parte de um provedor externo de dados

As responsabilidades de adequação impostos pelo GDPR (e pela LGPD) se referem também à obrigação legal de que um *recebedor* de dados de um provedor externo assegure-se do *compliance* deste fornecedor à disciplina vigente de proteção de dados pessoais com relação aos dados recebidos (GAL e AVIV, 2020, p. 18). Isso gera custos, dos quais alguns são reputados diretos, e outros indiretos. Custos diretos para assegurar que um fornecedor de dados esteja adequado à disciplina de proteção de dados pessoais incluem revisar o conjunto de dados recebido (quanto aos tipos de dados que contém, por exemplo), a forma com que foi coletado, a base legal utilizada para tanto etc. Custos indiretos podem incluir a possibilidade de perder a capacidade de utilizar os dados adquiridos por falta de adequação legal, ou os custos de discernir entre dados pessoais de outros tipos de dados. Embora esses custos variem muito com relação à atividade desempenhada, observa-se positivamente que não são relevantes no contexto de coleta interna de dados ou obtenção de conhecimento baseado em dados, e, desta forma, conclui-se que, quanto mais difícil ou custoso seja assegurar o *compliance* de um provedor de dados, mais forte se afigura o incentivo para que sejam adotados estes dois modelos de negócios para obtenção dos dados relevantes para a atividade (GAL e AVIV, 2020, pp. 18-9). Gal e Aviv (2020, p. 190) observam também um efeito indireto, sobre a estrutura de mercado, desta necessidade de assegurar o *compliance* de um provedor externo: geralmente, incluem-se clausulas contratuais quando do fornecimento dos dados com a finalidade de assegurar segurança jurídica, o que, sob o ponto de vista do fornecedor de dados, configura situação em que os incentivos para que atue no mercado são limitados pelos custos de se comprovar o *compliance*, de viabilizar o monitoramento constante de suas atividades, e de sanções impostas pelo recebedor dos dados em caso de alguma falha.

4.3.3 Garantia de que o acesso compartilhado ou que dados compartilhados sejam utilizados em conformidade com a disciplina de proteção de dados pessoais.

Para além de obrigações impostas àqueles agentes que recebem dados pessoais para realizar suas operações, o GDPR impõe a fornecedores de dados a responsabilidade de garantir a conformidade dos agentes que receberão dados pessoais, i.e., seus clientes, com obrigações eventualmente pré-estabelecidas para com os titulares de tais dados. No caso do GDPR, se deduz estas obrigações da responsabilidade do controlador de dados para com o titular de garantir que a atividade realizada com seus dados pessoais o serão de acordo com o seu consentimento (quando esta for a base legal aplicável), incluindo o direito unilateral do titular de solicitar o apagamento dos dados, a qualquer tempo (GAL e AVIV, 2020, p. 19), o que, conforme as normas jurídicas aplicáveis, deverá ser atendido pelo responsável pela obtenção do consentimento com relação a todos os usos ulteriores, decorrentes deste consentimento, i.e., base legal, que assegurou a licitude da coleta dos dados na origem. Assim, uma empresa que detenha dados pessoais para fornecer a terceiros, e que esteja vinculada ao consentimento de um titular destes dados, ao fornecê-los a terceiro agente de tratamento de dados no contexto de atividade econômica legítima, deverá assegurar que os dados serão utilizados conforme o consentimento dado pelo titular, e poderá ser objeto de solicitação pelo titular, amparado pelo GDPR, de apagamento dos dados fornecidos, devendo ter meios de garantir que aquele que recebeu os dados o faça.

Deter os meios de garantir que o recebedor dos dados os apague no caso de solicitação por parte do titular é apenas uma faceta da situação delineada, que consiste na responsabilidade da pessoa que realiza atividade econômica legítima relacionada ao fornecimento de dados pessoais para com relação a *todos os usos ulteriores* destes dados, inclusive após a sua transferência de forma legítima e consentida pelo titular. Ilustrativamente com relação ao crescente risco em atividades em geral que envolvem compartilhamento de dados, a Corte de Justiça da União Europeia recentemente afirmou que o operador de um *website* é responsável, em conjunto com os provedores de tecnologias empregadas em seu *website*, por eventual descumprimento do GDPR pelos últimos, mesmo que seja estabelecido que o *website* e seu operador não podem factualmente, i.e., devido a restrições tecnológicas, controlar quais dados pessoais são transmitidos para ou tratados pelo referido provedor de tecnologias (GAL e AVIV, 2020, p. 20). Não é difícil perceber o potencial deste tipo de norma de propiciar concentração

de mercado, ao passo que desestimula novos entrantes. A única ficção que justificaria esta postura seria a crença na *eliminação* das atividades com dados pessoais.

Em decorrência do regime jurídico da Lei Geral de Proteção de Dados Pessoais brasileira, podemos deduzir a ocorrência das mesmas situações jurídicas, uma vez que o plexo normativo é similar nos aspectos delineados.

Gal e Aviv (2020, p. 19), diante de incerteza jurídica ainda predominante neste contexto no continente europeu, inclusive pelas escassas oportunidades para tribunais enfrentarem controvérsias a ele relacionadas, argumentam que, no mínimo, o fornecedor de dados deverá poder razoavelmente considerar que o recebedor dos dados pessoais que fornece os utilizará em conformidade com o GDPR com relação a quaisquer destes dados. O compartilhamento de dados, assim, aumenta o grau de risco a que está sujeito aquele que os fornece, pois este pode não ter controle sobre o recebedor dos dados, configurando um cenário de *risco moral* e grau ao menos moderado de incerteza. Para lidar com isso, deverão incidir custos relacionados a ações como constante monitoramento e realização de auditorias sobre a pessoa que recebe os dados pessoais, que deverão ser mais intensas conforme o grau de risco identificado. O grau de risco será composto por vários fatores. Destaca-se, com relação ao compartilhamento de dados, o incremento de risco que advém da eventual *combinação* de conjuntos de dados, i.e., do fornecedor e do recebedor dos dados, que aumenta consideravelmente a sua sensibilidade[12]. Destaca-se, ainda, o incremento na incerteza decorrente desta combinação nas hipóteses em que apenas uma das partes está exposta à combinação dos conjuntos, enquanto a outra parte apenas conhece o próprio conjunto dos dados (o que poderia ocorrer no contexto de uma empresa *fornecedora de dados pessoais* para uma pessoa que os adquire ou por outra forma contratual os recebe para realizar suas operações que necessitam de dados pessoais).

4.3.4 Obrigações relacionadas à gestão de dados.

O GDPR, como a LGPD, impõe deveres àqueles responsáveis por dados pessoais durante todo o período que se convencionou chamar de *ciclo de vida* do dado. Por exemplo, deve-se assegurar que os dados sejam acessados apenas nos casos em que for necessário, por pessoas autorizadas, e de forma a ser passível a identificação de responsáveis por incidentes;

[12] Cf. Solove, 2006; 2012; 2020.

deve-se manter dados pessoais pelo menor prazo possível para o atingimento da finalidade destacada, e no menor grau necessário. Por sua vez, pode ser preciso impor controles e limitações físicas e tecnológicas para atingir este grau de controle, assim como treinar funcionários continuamente e estabelecer clausulas contatuais hábeis a delimitar responsabilidades internamente. Assim, o cumprimento dessas obrigações com grau razoável de segurança jurídica pode envolver altos custos com reorganização interna da companhia, dos quais se destaca a necessidade de se monitorar o fluxo de dados pessoais individualmente considerados na empresa, de forma a gerir a sua acessibilidade, as formas de seu uso e assegurar a manutenção de sua regularidade.

Tais processos de gerenciamento de dados cuja adoção é imposta pela estrutura normativa são característicos da estrutura de economias de escala, o que representa a criação de uma vantagem competitiva para grandes agentes de tratamento de dados com relação aos menores ou aos que desejem entrar no mercado. Estas empresas com grande abrangência de suas atividades têm possibilidades como compartilhar ou reduzir custos através de um *data pool* ou uma fusão, ou mesmo evitar os custos decorrentes desta responsabilidade através da aquisição direta de conhecimento baseado em dados pessoais.

Especialmente no continente europeu, tem sido cada vez mais presentes soluções de mercado voltadas a facilitar este controle, dos quais se destaca o uso das Plataformas de Gerenciamento de Consentimento (na sigla em inglês, CMP), que têm sido adotadas por agentes de tratamento de dados sob a jurisdição do GDPR mesmo em casos em que não é imposta pela norma (GAL e AVIV, 2020, p. 21). As CMPs rastreiam a transferem informações relativas ao consentimento fornecido (ou retirado) por cada titular de dados para quaisquer propósitos, e permitem tanto àquele que fornece, quanto ao que recebe dados pessoais, verificar que o consentimento necessário para as operações envolvidas com os dados em questão foi concedido. Estas ferramentas reduzem alguns dos riscos mencionados anteriormente (GAL e AVIV, 2020, p. 21).

4.3.5 Obrigações legais relacionadas ao tamanho e ao tipo da atividade com dados pessoais.

No caso do GDPR, Gal e Aviv (2020) apontam que as características da base de dados podem afetar a imposição de restrições adicionais a atividades de tratamento que se lhe refiram.

Por um lado, quanto maior o volume, variedade, veracidade e velocidade de uma base de dados, mais valioso será o conhecimento dela extraído e, portanto, maior valor comercial terá a referida base de dados. Por outro lado, quanto maior a chance de que dados possam ser utilizados para extrair informações sensíveis acerca de um indivíduo ou que possam de alguma forma identifica-lo, mais estritos são os requisitos impostos pelo GDPR com relação à atividade de tratamento de dados pessoais.

No caso da LGPD, este aspecto é, primeiramente, extraído das disposições gerais do diploma, que trazem um rol de fundamentos e objetivos da disciplina, que se referem a direitos de natureza constitucional cujos âmbitos de incidência da respectiva proteção, eventualmente, se autodelimitarão. Ilustrativamente, com relação às bases legais alternativas para viabilizar o tratamento de dados, a LGPD apresenta a possibilidade de adoção do legítimo interesse do controlador, especificando que não poderá se sobrepor a direitos dos titulares que dependam do resguardo de seus dados do uso pretendido. Neste caso, o controlador é obrigado também a realizar uma avaliação e um relatório acerca do potencial impacto da atividade sobre os direitos os titulares e as medidas de mitigação adotadas. Outro exemplo é a criação, pela LGPD (a exemplo do GDPR), da categoria de dados pessoais *sensíveis*, sobre os quais recai maior proteção legal, como, por exemplo, um rol diferenciado e mais limitado de bases legais possíveis – nas quais não se inclui o interesse legítimo do controlador, pelo que se pode deduzir que, sob o regime da LGPD, não há atividade de tratamento de dados pessoais sensíveis que se justifique por legítimo interesse do controlador.

No continente europeu, este contexto exibe uma característica que, até o momento, não foi trazida para o Brasil quando da instauração do regime jurídico decorrente da LGPD: a imposição de determinadas obrigações apenas para agentes de tratamento de dados com determinado tamanho ou natureza da operação. O *Data Protection Officer*, por exemplo, não precisa ser contratado por todas as companhias, e empresas com menos de 250 funcionários não são obrigadas a manter registros acerca de atividades com dados pessoais, exceto caso estas atividades não sejam incidentais ou envolvam informações sensíveis. No Brasil, a Autoridade Nacional de Proteção de Dados Pessoais tem atuado para orientar agentes de tratamento de dados, a partir, por exemplo, da publicação do Guia de Segurança da Informação para Agentes de Tratamento de Pequeno Porte, em 2021. A edição de normas específicas para diferentes grupos de agentes, incluindo os de pequeno porte, ainda é esperada.

Por outro lado, ainda quando há gradação, é notável que a magnitude destas obrigações cresce apenas até o ponto em que se atinge o maior grau reconhecido pela norma em relação ao tamanho ou sensibilidade do conjunto de dados. Após este ponto, os custos marginais para se manter adequado às obrigações legais decrescem em virtude das características de economias de escala. Neste ponto, a empresas de grande porte é atribuída vantagem competitiva pelo regime jurídico de proteção de dados pessoais, com relação a competidores de médio ou de pequeno porte (GAL e AVIV, 2020, p. 22). No caso do Brasil, este desequilíbrio é mais grave do que no continente europeu, em que empresas de *pequeno* porte e sem atividades relevantes relacionadas a dados pessoais são eximidas de algumas das obrigações legais referidas.

Em outras palavras, quanto maior a empresa, menores os seus custos de *compliance per datum*, relativamente a empresas menores que precisem cumprir com requisitos similares.

4.4 Efeitos do GDPR sobre escolhas para obtenção de dados pessoais

Nesta seção, exploraremos como, em conjunto, os fatores delineados, decorrentes da disciplina jurídica de proteção de dados pessoais, afetam a escolha de empresas que necessitam de dados pessoais em suas operações com relação ao modelo de negócio adotado para a sua obtenção, a partir da análise de Gal e Aviv (2020).

Consideremos a situação hipotética de a Empresa A necessitar, para o desempenho de suas operações legítimas, de dados pessoais, que podem alternativamente ser coletados internamente, ou obtidos de uma empresa externa (Empresa B). Ausentes barreiras tecnológicas, estratégicas e jurídicas à obtenção e compartilhamento de dados pessoais, a escolha será baseada nos limites financeiros, ou seja, será baseada no custo relativo de se coletar os dados internamente (coleta primária de dados, ou *first-party data*), ou obtê-los de uma fonte externa (coleta secundária de dados, ou *third-party data*) (GAL e AVIV, 2020, p. 22).

Na presença de limites jurídicos impostos pela disciplina de proteção de dados pessoais, outros aspectos devem ser sopesados. Adquirir dados da Empresa B traz diversos custos e obstáculos para a Empresa A, correspondentes aos esforços que deverão ser destinados a assegurar que há uma base legal adequada para a utilização dos dados adquiridos, e que as atividades da Empresa B em geral com relação aos dados compartilhados ocorrem em conformidade com a legislação. Ao adquirir um conjunto de dados de fonte externa, a empresa

A poderá, ainda, ter de adequar o ciclo de vida dos dados internamente às suas atividades para assegurar que se coadunam com os dados compartilhados. A Empresa B também incorre em custos, especialmente os relativos a assegurar que a Empresa A, recebedora dos dados, os utilizará apenas dentro dos limites legais, e atenderá a eventuais solicitações de titulares com relação aos direitos conferidos pela lei. Quanto maior for a perda de controle sobre os dados por parte de seu detentor original, maiores são os riscos de não adequação legal, e maiores os custos relativos (GAL e AVIV, 2020, p. 23).

Estes custos e obstáculos limitam os incentivos a Empresa A para que adquira dados da Empresa B, ou para que estas engajem-se na formação de um *data pool* conjunto. Pode, ainda, resultar em menor competição, por parte de controladores de dados, com relação à oferta de dados pessoais, considerados os custos envolvidos na obtenção dos dados. Por outro lado, se os custos de assegurar ou demonstrar o *compliance* ao coletar dados para compartilhamento forem muito altos, menos competidores entrarão neste mercado também por esta razão, aumentando a concentração. Observe-se que as barreiras legais existentes reduzem as possibilidades de que A adquira os dados de que necessita, e desta forma aumenta os incentivos para que a coleta de dados seja feita internamente, ainda que esta opção seja reputada ineficiente sob os demais aspectos (tecnológico, econômico, estratégico).

Além de adquirir os dados de que necessita, a Empresa A poderia optar por realizar uma fusão com a Empresa B. Ausentes limites legais, os dados da empresa B poderiam ser utilizados internamente pela nova empresa após a fusão. No entanto, conforme discutido anteriormente, não são claros pela legislação ou pela sua interpretação por autoridades competentes, as condições e os limites para a utilização de dados pessoais após uma fusão. No mínimo, é de se esperar que limitações incidirão com relação às finalidades originais para as quais os dados foram coletados, que deverão ser mantidas, e ainda com relação à observância dos direitos dos titulares, que possivelmente aquiesceram com o uso dos dados em contexto diverso. Assim, quando de uma fusão, deverão ocorrer os custos relacionados a assegurar o *compliance* da empresa com que se deseja fundir (para obter e utilizar a sua base de dados) e a legitimidade da base legal que embasou a coleta e o uso dos dados. Gal e Aviv (2020, p. 23), ao realizar entrevistas com executivos relacionadas ao impacto do GDPR em seus negócios, receberam de mais da metade dos entrevistados a resposta de que estes já haviam trabalhado em uma potencial fusão de negócios que não foi levada a cabo em razão de preocupações associadas ao GDPR.

Assim, além de escolhas relativas ao modelo de negócios adotado, o regime jurídico de proteção de dados pessoais afeta considerações relacionadas a com quem fundir negócios, de quem adquirir dados pessoais e de com quem compartilhar dados pessoais ou manter bases de dados compartilhadas. Em regra, a empresa A terá incentivos mais fortes para se fundir com ou para comprar dados de uma empresa cujo *compliance* seja mais rápido e menos custosamente verificável, o que afeta a competitividade ao criar uma vantagem comparativa para fornecedores de dados já existentes e que já tenham sido escrutinizados com relação ao *compliance* às normas aplicáveis de proteção de dados pessoais. Deverá, ainda, sob esta ótica, ter maiores incentivos para lidar com empresas com reputação estabelecida no mercado, ou com um grande fornecedor de dados pessoais, ao invés de várias empresas menores (GAL e AVIV, 2020, p. 24).

A dinâmica relatada aumenta os custos do compartilhamento de dados, seja por meio de uma transação (compra) única, ou por meio de uma fusão ou *joint venture*, e desta forma reforça os incentivos para a realização de coleta de dados internamente, mesmo que uma empresa externa possa fazê-lo de forma mais eficiente, ou já tenha investido recursos previamente para coletar dados. Em casos extremos, e naqueles em que os dados pessoais não sejam essenciais para a atividade desempenhada, os custos do *compliance* podem também limitar a própria coleta interna de dados. Neste sentido, Gal e Aviv (2020, p. 24) observaram que, em pesquisa realizada com profissionais da área de proteção de dados, 20% afirmaram que a adequação completa ao GDPR é impossível, o que resulta no referido processo de *balcanização*. Neste contexto, parece fundamental questionar os efeitos que as normas têm exercido sobre atividades econômicas legítimas que dependem de dados pessoais, ou sobre pessoas com potencial de prestar serviços mais valiosos para a sociedade por meio de operações que os envolvam.

As limitações relacionadas ao compartilhamento de dados pessoais podem também criar ou reforçar incentivos para que sejam adotados modelos de negócios sob os quais uma empresa pode controlar todos os produtos e serviços no ecossistema relevante (GAL e AVIV, 2020, p. 24). Há vários exemplos. Um deles se refere à estratégia adotada pela Verizon ao adquirir a sua cadeia de fornecimento de dados (AOL, Yahoo! e outras) para melhor controlar e utilizar dados de consumidores voltados a anúncios personalizados e melhoras nos serviços prestados. No mesmo sentido, o Google adquiriu a DoubleClick e outros ativos e tecnologias similares envolvidos no seu ecossistema de fluxo de dados. Observa-se que quanto maiores os obstáculos criados pelo regime jurídico de proteção de dados pessoais ao compartilhamento voluntário de dados entre diferentes pessoas, mais fortes serão os incentivos para este tipo de expansão interna (GAL e AVIV, 2020, p. 24).

O regime do GDPR também cria vantagens para empresas de grande porte, reduzindo a potencial competitividade no mercado sobre o qual é aplicável. Gal e Aviv (2020) identificaram quatro razões principais para esta ocorrência.

Primeiro, como observado, empresas de grande porte se aproveitam de características inerentes a uma economia de escala no contexto de coleta e gerenciamento de dados que seja conforme à disciplina jurídica de proteção de dados pessoais. No caso do GDPR, o reconhecimento deste efeito leva à imposição de obrigações mais lenientes a empresas de pequeno e médio portes (GAL e AVIV, 2020, p. 25). No caso da LGPD, embora haja a previsão de que a Autoridade Nacional de Proteção de Dados Pessoais poderá criar critérios diferenciados tendo por base parâmetros desta natureza, até o momento não houve regramento neste sentido, o que significa que as obrigações impostas pela lei, em princípio, afetam igualmente empresas de todos os portes, e sem diferenciação quanto à posição da atividade de tratamento de dados nas operações da companhia, e.g., se incidental ou se objeto de sua atividade principal. Gal e Aviv (2020, p. 24) apontam, porém, que em algumas situações empresas de menor porte podem não se beneficiar desta maior leniência, pois um menor padrão de adequação pode colocá-las em posição de desvantagem com relação àquelas com melhores padrões, diante das exigências legais para a contratante, levando-as a serem preteridas caso não suportem os custos mais elevados da adequação a padrões mais elevados. Uma empresa A que deseje fundir o conjunto de dados da empresa B a seu próprio poderá estar mais inclinada a fazê-lo com relação a conjuntos de dados que já estejam adequados a um padrão de conformidade à norma mais elevado, ainda que a outra parte não tivesse a obrigação legal de seguir padrões mais rigorosos.

A segunda razão se refere ao alto grau de incerteza decorrente de algumas das disposições do GDPR, que podem impor custos mais altos a participantes menores. Alguns exemplos recorrentes são as incertezas relativas aos limites para a delimitação da licitude de uma atividade de tratamento de dados, e também à abrangência territorial exata do GDPR. Outro exemplo desta incerteza é encontrado com relação ao direito à portabilidade, assegurado tanto pelo GDPR, quanto pela LGPD. O direito à portabilidade reconhece que um titular poderá ter acesso aos seus dados, ou tê-los transferidos a terceiro, em formato estruturado e legível por máquina. O objetivo primevo, em linhas simples, é assegurar que o titular poderá trocar de provedor de serviços quando desejar, inclusive levando ao concorrente os seus dados pessoais em poder do provedor atual. O direito à portabilidade, assim, visa a *remover*, ou enfraquecer, o incentivo negativo ao titular para trocar de provedor de serviços quando deseje em virtude de

perda na qualidade do serviço prestado que poderia advir do fato de o novo prestador não os possuir. Embora esta garantia legal afete também empresas de grande porte, duas razões permitem afirmar que estas gozam de vantagem competitiva neste cenário com relação aos seus concorrentes de menor porte. O primeiro se relaciona ao fato de que, em economias de escala, quaisquer custos relacionados à manutenção de padrão quanto os dados pessoais serão distribuídos por todo o conjunto, desta forma sendo mais baixos *per datum* quanto maior seja a atividade de tratamento de dados. Em segundo lugar, na falta de facilitação por parte do governo – como, por exemplo, o estabelecimento de padrões a serem observados quanto ao formato e outras características técnicas dos dados objeto de portabilidade –, os padrões de operabilidade e interoperabilidade tendem a ser determinados pelos grandes *players* do ramo, que o farão conforme suas próprias preferências, inclusive aumentando os custos de *compliance* para seus concorrentes de menor porte. O vácuo na delimitação do conteúdo da norma, assim, pode levar ao artificial erguimento de barreiras tecnológicas e estratégicas à interoperabilidade dos dados, o que representa outro fator de favorecimento à concentração de mercado.

A terceira razão também se relaciona à incerteza com relação à adequada interpretação de alguns dos dispositivos do GDPR, que podem conduzir ao uso estratégico de sua interpretação como uma tática ofensiva. Após a introdução do GDPR, muitas grandes empresas de tecnologias limitaram práticas anteriormente empregadas de compartilhamento de dados, sob a justificativa de adequação à norma. No entanto, em muitos casos, observou-se que as limitações iam além do que poderia ser razoavelmente deduzido das disposições do GDPR, sugerindo o seu uso estratégico. O Google, por exemplo, limitou sob esta justificativa operações de transferência de dados relativas aos lances de interessados em arrematar espaços de publicidade. A este respeito, Geradin e Katsifis argumentaram que estas limitações poderiam representar a "morte" deste mercado, pois editores perderiam a possibilidade de medir e avaliar o incremento de valor relativo ao marketing, com relação àquele realizado nos canais controlados pelo Google, ao passo que o GDPR não justificava a conduta, embora esta fosse a alegação da empresa (GAL e AVIV, 2020, p. 26). A norma, assim, foi instrumentalizada para que o Google retirasse a publicidade de informações que poderiam favorecer a entrada de novos participantes no mercado, ou seja, para manter a sua fatia de mercado.

O quarto fator se relaciona ao fato de que o tamanho de uma empresa, e sua reputação no mercado, podem afetar a conduta de titulares de dados, embora o GDPR e a LGPD não sejam voltados afetar diretamente a sua conduta. É razoável supor que as pessoas em geral considerarão que os detentores de grandes conjuntos de dados estarão sujeitos a maior

escrutínio de suas atividades por autoridades do que os menores competidores, e que empresas com conjuntos de dados de alto de valor, ou que sejam muito relevantes para o seu modelo de negócios, tem muito mais a perder se não se adequarem à LGPD do que empresas com conjuntos de dados de baixo valor, ou que são relevantes para as operações da empresa apenas incidentalmente (GAL e AVIV, 2020, p. 26). Sob outro aspecto, também é de se esperar que grandes empresas de tecnologia tenham maiores incentivos e capacidade financeira de proteger seus conjuntos de dados de ameaças externas. Esses fatores representam um potencial favorecimento a estas empresas em detrimento de menores participantes do mercado, quando titulares de dados são instados a fornecer seus dados para acesso a um serviço ou produto que seria acessível a partir de múltiplos provedores. Este maior incentivo também pode ser observado no fato de que usuários podem preferir limitar o número de empresas para as quais fornece consentimento. Estes fatores também favorecem a concentração nos mercados relacionados a dados pessoais (GAL e AVIV, 2020, p. 27), contrariando o objetivo de favorecer a desconcentração de poder.

Além dos casos explorados, em que os dados em questão poderiam ser coletados pelas Empresas A ou B de forma intercambiável, tomemos a situação em que os conjuntos de dados de cada uma, ao serem combinados, criam sinergias de dados que aumentam a qualidade das informações e do conhecimento que poderiam ser extraídos de cada um considerado isoladamente. Estas sinergias podem ser operacionalizadas pela aquisição dos dados, pela criação de uma *joint venture*, ou por uma fusão entre as empresas. O regime jurídico de proteção de dados vigente não estabelece regras diferenciadas na ocorrência destas sinergias relativamente a quando não estão presentes. É notável, porém, que estas sinergias podem envolver custos e riscos adicionais. Por exemplo, a combinação de um conjunto de dados interno com outro obtido de fonte externa, que são regidos por políticas de privacidade diversas, pode significar a necessidade de que a empresa realizando operações com o conjunto união dos dados redesenhe sua arquitetura interna de fluxo de dados e suas políticas internas sobre dados pessoais. As sinergias de dados também podem afetar a extensão da aplicabilidade das normas jurídicas, quando, por exemplo, a combinação dos dados levar a ocorrência, no conjunto, de dados reputados sensíveis (GAL e AVIV, 2020, p. 28)

Estes resultados são consistentes com os encontrados por outros estudos recentes quanto ao mundo real e a partir de dados. Johnson et. al. (2020; 2020b), observaram em extensa e minuciosa pesquisa empírica que o GDPR opera o efeito de restringir a utilização de provedores de tecnologias da *web* por detentores de *websites*, ao passo que promove maior concentração

de mercado entre provedores de tecnologias da *web*. Johnson et. al (2020b) notam que sua análise acerca do papel do GDPR quanto à concentração de mercado contribui para a literatura sobre a matéria especialmente em razão de ser capaz de distinguir entre o impacto do GDPR sobre a atividade econômica real, e os impactos que recaem tão somente sobre a forma de se registrar resultados econômicos, i.e., modificações no comportamento do titular de dados com relação à aquiescência com atividade de tratamento de dados pessoais. A sua pesquisa, a partir de metodologia explicitada no estudo, é capaz de identificar a redução no resultado econômico real como efeito das disposições do GDPR. Vale dizer que estudos anteriores foram capazes de estabelecer relações similares: por exemplo, foi documentado que a restrição sobre a propaganda de cigarros levou a maior concentração de mercado no respectivo setor (ECKARD, Jr., 1991; GALLET, 1999; CLARK, 2007), bem como no mercado de álcool (SASS e SAURMAN, 1995), reforçando a importância de se enxergar a regulação jurídica – existente ou potencial – a partir de sua interação com outros aspectos existentes no mundo, i.e., reguladores (LESSIG, 1999), de forma a viabilizar a sua inserção no contexto que intende regular de forma adequada e apta a atingir os objetivos almejados, ao passo que se minimiza os resultados indesejados, ou ao menos se viabiliza um caminho para se identificar estes efeitos deletérios indesejados, de modo a evitar a chamada *opacidade causal* tão comum na intervenção em sistemas de complexidade organizada.

CONSIDERAÇÕES FINAIS

A virtualização, e sobretudo a sua intensificação a partir do processo de digitalização da cultura humana, pela centralidade e globalidade que adquire na sociedade, representa traço marcante de nosso tempo (NELLIS, 1991; 2014), e com isso lidar e dominar a fenomenologia inerente a este processo torna-se fundamental para a permanência da utilidade de qualquer instituição ou campo do conhecimento, e com o direito não é diferente. Isto é reforçado pela observação de práticas culturais contemporâneas, em que a participação de pessoas em processos de coleta, transformação e difusão de informações e obras culturais, através das fronteiras jurídicas de "pessoais" (no contexto da disciplina de proteção de dados pessoais) ou "obras protegidas" (no contexto da Lei de Direito Autoral), adquirem papel marcante e central nas vivências individual e coletiva. A inobservância de critérios objetivos para a determinação das fronteiras dessas categorias pode levar a uma utilização instrumental do direito para

interesses que desbordam do intuito do ordenamento jurídico-constitucional, ou mesmo que lhe firam. Por outro lado, ao interferir em qualquer dinâmica social, o direito corre o risco de causar diversos danos, que apenas são perceptíveis caso se estruture o conhecimento em torno da disciplina e dos fenômenos regulados e a ele relacionados de forma adequada.

Se, no advento da *internet,* as aspirações para a rede se associaram precipuamente àquelas de cunho libertário, a partir sobretudo dos anos 2000 houve um maior alargamento do debate, com o fortalecimento de posições favoráveis à ocupação de um espaço regulatório da rede pelo Estado, embora não ligados a correntes de pensamento tradicionalmente estatistas ou autoritários. Esta regulação seria voltada a garantir a permanência da rede como um espaço público, livre e aberto para a inovação. Ao mesmo tempo, acomodavam-se pretensões de alcance das jurisdições nacionais a atos ilícitos cometidos na rede. O desejo de que os estados nacionais, em alguma medida, pudessem alcançar atividades ocorridas na *internet*, foi um catalisador de regulamentações em torno da arquitetura e atividades na rede. Em associação com o contexto delineado de regulação da rede para manutenção das liberdades, foi comum o endereçamento de preocupações em torno de concentração e outras falhas de mercado, especialmente quando estas pudessem levar à censura, controle indevido ou abuso de poder econômico.

O crescimento das chamadas *big techs* apenas fez e faz reforçar este debate, uma vez que a falha do estado em regular efetivamente o ambiente virtual tem permitido a atores econômicos tomar frente e ocupar cada vez mais espaços, inclusive sob o ponto de vista regulatório. Se, por um lado, não se busca uma expansão da atividade regulatória para além do necessário, é preciso observar que o vácuo de atuação do estado, seja por inação ou ação inefetiva, é eventualmente ocupado, nomeadamente na ocorrência de desequilíbrios de poder. Esta ocupação pode se dar de várias formas, como por meio de abuso de poder econômico, ou por atores mal-intencionados imbuídos de funções estatais que se veem diante de uma estrutura jurídica de fácil manipulação, pelo que se revela a importância de estrutura normativa sólida sobre o tema, embasada em aportes teóricos consistentes e resistentes à manipulação. O vácuo de atuação do Estado em determinada dinâmica social pode decorrer não apenas de inação em sentido estrito ou da inexistência de normas jurídicas a seu respeito; frequentemente, pode ocorrer de determinado regime jurídico perder relevância para a sociedade e atrair o desprezo e descrédito públicos, ou de impor obrigações que são cumpridas de forma aparente ou balcanizada, o que representa ultimamente um vazio de atuação estatal na manutenção de

âmbitos de liberdade constitucionalmente protegidos, especialmente diante o caráter normativo da arquitetura da rede e da concentração de poder econômico marcante no setor de tecnologia.

No capítulo I, introduzimos os pressupostos em que se calcam as abordagens ora propostas, especialmente a importância da compreensão do objeto que se visa a regular e das relações que o direito pode com ele estabelecer. Levamos em especial consideração, ainda, a relação entre fundamentos e objetivos, e as formas com que a abordagem fenomenológica pode auxiliar no estabelecimento de relações fáticas viáveis, efetivas e em consonância com os objetivos fundamentais que se persegue nos campos regulatório e judicial. Tratamos, ainda, da forma com que metáforas podem e são efetivamente empregadas para a transmissão de conceitos acerca de proteção de dados pessoais e privacidade. Destacamos que são mecanismos sofisticados e hábeis a estabelecer a visão acerca de um objeto ou problema e defini-lo, ao passo que restringe este objeto e as suas faces visíveis. No âmbito da privacidade e proteção de dados pessoais, vimos que diversas metáforas têm sido empregadas para estabelecer a compreensão acerca do tema, com diferentes enfoques e amplitudes.

No capítulo II, foi trazido um seletivo e breve histórico do tratamento jurídico declinado à tutela da informação, que, sob as lentes da fenomenologia, foi colocado lado a lado ao tipo de materialidade que se visa a tutelar. Foi possível observar a progressiva abertura do conceito de dado pessoal, especialmente a partir do momento em que se passou a tutelar a formação de perfis de informações sobre pessoas, o que abriu caminho para a associação do futuro conceito de dado pessoal a qualquer tipo de informação com o potencial de identificar um indivíduo. Vimos que este aspecto foi decisivo para a associação definitiva da tutela da informação àquela de privacidade, esta entendida de forma a abranger o direito à autodeterminação informacional. Este conceito ganhou especial notoriedade a partir da importância e abrangência crescentes da disciplina de proteção de dados pessoais no continente europeu, notadamente a partir do GDPR.

No capítulo III, tratamos do controle individual sobre a informação pessoal, inicialmente a partir de uma descrição dos maiores problemas relacionados com a compreensão e gestão da informação pessoal e da forma com que a disponibilização de informação pode auxiliar nestes processos. Aprofundamos este estudo a partir da abordagem de diversos estágios em que a informação pode ser recebida e empregada por um indivíduo, a partir de estreita correlação entre a sua utilidade ante a materialidade em que é instada a atuar. Argumentamos, ainda, pelo emprego de um tipo de compreensão multimodal do discurso presente em interfaces

digitais, de forma a potencializar a utilidade da informação tendente a tutelar dados pessoais e incrementar as possibilidades de autonomia diante do crescente uso de tais interfaces.

No capítulo IV, tratamos do papel dos dados na economia contemporânea, a forma com que são empregados e a relação destes aspectos com alguns cânones da disciplina de proteção de dados pessoais europeia, notadamente a existência de algumas antíteses que prejudicam o sopesamento de interesses contido na disciplina. Tratamos de mercados de dados pessoais, a forma com que são regulados e os efeitos decorrentes da interação de tais mercados com a regulação sobre dados pessoais. Observamos que alguns dos interesses que deveriam ser promovidos pela disciplina de proteção de dados pessoais, como a promoção da inovação e a desconcentração de poder, parecem sofrer efeitos deletérios em decorrência da regulação, para além da ineficiência em prover controle individual sobre a informação.

O presente estudo, com isso, se propôs a colocar em perspectiva a disciplina de proteção de dados pessoais conforme vem sendo erguida no Brasil a partir de seus princípios fundantes, confrontando-a com a facticidade sobre a qual deve atuar e com os problemas que é vocacionada a endereçar, com o objetivo de fomentar um debate que torne possível que a norma atinja resultados significativos, e especialmente não culmine em resultados negativos como os que têm sido constatados no continente europeu. Se é verdade que o enfrentamento destes problemas é fundamental para a sociedade, também o parece ser o fato de que eles não têm sido, de fato, enfrentados pela abordagem que o direito lhes tem concedido – e podem mesmo estar sendo exacerbados ou protegidos por esta abordagem. Por outro lado, o controle sobre o conhecimento passível de ser extraído de conjuntos de dados tem papel estruturante e transformador na sociedade que desborda do âmbito de proteção do direito à privacidade, ainda quando esta é concebida sob a ótica do papel no desenvolvimento autônomo do indivíduo. Com isso, uma normativa inadequada não apenas falha em resguardar suficientemente o âmbito de proteção relativo ao direito à privacidade e à autodeterminação informacional, como favorece a invasão do âmbito de proteção de outros direitos fundamentais ao não prover parâmetros para a resolução de conflitos e mútua delimitação do âmbito de proteção dos direitos fundamentais envolvidos, potencializando a concentração de poder em um mercado naturalmente falho dado o estado da arte da tecnologia e do conhecimento médio.

O denso conceito de *partido* pode ser utilizado para sintetizar as funções das discussões empreendidas no capítulo 1, e que serve como norte para a leitura deste trabalho: denota aquilo que, por um lado, representa o fundamento de algo, i.e., de um projeto, de uma norma, de uma

decisão etc. Por outro, denota o próprio *objetivo* deste algo. Possivelmente neste sentido, argumentou-se no campo de decisões judiciais e de políticas públicas que justiça e racionalidade apenas adquirem significado no contexto de nossos objetivos fundamentais. O *partido*, pela sua importância elementar quanto ao objeto a que se refere, serve ainda como parâmetro para a mensuração do atingimento dos objetivos que enuncia, e, ainda mais importantemente, como parâmetro de balanceamento e ponderação entre conflitos que surjam *no âmbito do algo,* i.e., do projeto, da norma, a que se refere, e, desta forma, solucionar conflitos existentes ao longo do processo de maneira coerente. Afinal, caso, no processo de realizar algo, tenhamos uma dúvida entre duas ou mais escolhas que não possa ser resolvida meramente pela técnica, esta escolha, para que tenha sentido no âmbito do que é feito, deverá ser realizada com base no próprio partido, ou seja, com base naquilo que se justifica quanto aos *fundamentos* e aos *objetivos* por ele enunciados, e que são a própria razão de ser e a essência deste algo de que estamos falando, i.e., um projeto, um Regulamento, um sistema de normas, uma política pública etc. Com isso, a inexistência de um *partido,* no sentido delineado, ou a confusão e indefinição a seu respeito, representam um grave problema caso se deseje, por meio deste algo, atingir objetivos consistentes. Afinal, a todo o momento um projeto ou um sistema normativo requer sejam tomadas decisões: no âmbito legislativo; em cada decisão administrativa ou judicial particular; na implementação de projetos de conformidade à lei etc. Caso os fundamentos não conduzam logicamente aos objetivos do sistema, eles poderão justificar decisões que não os persigam, ou mesmo que lhe firam. Por outro lado, aquilo que for realizado terá efeitos, ainda que indefinidos. O partido, assim, deve estar intrinsecamente conectado à materialidade sobre a qual o projeto atua.

Os inúmeros princípios e direitos de caráter fundamental elencados pela LGPD são representativos da importância da disciplina jurídica de proteção de dados para o resguardo das liberdades constitucionais e o asseguramento de um ambiente social, hoje intensamente mediado por tecnologias da informação, propício ao desenvolvimento individual e coletivo autônomos e significativos. Por outro lado, foi possível observar que os conceitos legais trazidos pela LGPD, especialmente aqueles relacionados à incidência da norma, são similares ao GDPR, e desta forma tendem a atrair problemas similares. Igualmente, a LGPD traz regime de obrigações e direitos bastante similar e que não apresenta graus de acordo com o tipo, o risco ou o benefício decorrente da atividade regulada, o que reforça a probabilidade de ocorrência dos problemas que de fato têm sido observados no continente europeu em decorrência do regime jurídico do GDPR. Além disso, a LGPD foi erguida sem o prévio estabelecimento do

tema da proteção de dados pessoais como um problema público, do que pode decorrer algumas de suas dificuldades em viabilizar a apropriação da norma para finalidades importantes, exacerbando o seu risco de captura.

A sua eficácia no atingimento de valores fundamentais e resultados significativos para a sociedade depende de sua integração com o direito e cultura nacionais de modo a contextualizar os direitos fundamentais e as afrontas ao seu âmbito de proteção a partir dos fenômenos contemporâneos, e, sobretudo, de sua integração com o conhecimento relativo aos processos de tratamento de dados pessoais. A partir da observação da disciplina de proteção de dados em panorama e de forma contextualizada, esperamos ter contribuído para o amadurecimento dos debates em torno dos direitos à privacidade e à proteção de dados pessoais, e para a sua interpretação no ordenamento jurídico de forma harmônica e que potencialize a sua aptidão para atingir resultados significativos.

Ao passo que a abertura metodológica e do objeto de pesquisa do presente trabalho representa uma limitação quanto à sua capacidade de empreender análises mais detidas sobre aspectos específicos da disciplina de proteção de dados pessoais, fornecendo respostas operacionalizáveis, frentes significativas de desenvolvimento e de pesquisa se mostram viáveis a partir das discussões empreendidas neste trabalho. Por exemplo, foi demonstrado que a conceituação dos termos mais elementares para a disciplina de proteção de dados, i.e., o conceito de informação e o conceito de dado pessoal, foram implementados em nosso ordenamento jurídico sem delimitação de suas fronteiras materiais e dos seus papeis, o que é um reflexo da forma com que já era trabalhado nas jurisdições de que os emprestamos. Isto compromete a efetividade da norma quanto aos seus objetivos propostos, e ao mesmo tempo ameaça o âmbito de proteção de outros direitos fundamentais, pelo que o consistente desenvolvimento destes conceitos é elementar para uma disciplina sólida e capaz de ser parametrizada e atingir objetivos fundamentais da sociedade. Outro ponto carente de desenvolvimento é o de *privacidade por design* e por padrão, especialmente a partir da compreensão das arquiteturas de escolha, do déficit epistemológico envolvido nas relações mediadas por interfaces e dispositivos e da multimodalidade dos diálogos neste âmbito. Seriam úteis estudos capazes de aplicar o *framework* proposto para a avaliação dos ambientes de interação de titulares de dados com relação aos padrões legais, inclusive com o emprego do proposto critério de articulabilidade quanto à clareza das informações transmitidas aos usuários.

REFERÊNCIAS

ACQUISTI, Alessandro; TAYLOR, Curtis; WAGMAN, Liad. The Economics of Privacy. 54 J. Economic Literature, pp. 442-478, 2016.

ALEXY, Robert. Theorie der Grundrechte. Suhrkamp, 1985.

ALEXY, Robert. Teoria da Argumentação Jurídica: A Teoria do Discurso Racional como Teoria da Fundamentação Jurídica. Introdução de Cláudia Maria Toledo. Rio de Janeiro: Forense, 2005.

ASHER, Nicholas; LASCARIDES, Alex. Logics of Conversation. Cambridge: Cambridge University Press, 2003.

ATASOY, O.; MOREWEDGE, C. K. Digital goods are valued less than physical goods. Journal of Consumer Research, 44(6), pp. 1343–1357, 2017.

BALKIN, J. M. Cultural Software: A Theory of Ideology. 1998.

BALL, Kirstie. Exposure: exploring the subject of surveillance. Information, Communication and Society, v. 12, n. 5, pp. 639-57, 2009.

BATEMAN, John A.; WILDFEUER, Janina. A multimodal discourse theory of visual narrative. Journal of Pragmatics 74, 2014.

BAUMAN, Zygmunt; LYON, David. Liquid Surveillance: A Conversation. Polity Press, 2013.

BEECHER-MONAS, Erica. Evaluating scientific evidence: an interdisciplinary framework for intellectual due process. New York: Cambridge University Press, 2017.

BENTHAM, Jeremy. Panopticon Or the Inspection House. London: T. Payne, 1791.

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

BIONI, Bruno Ricardo. Proteção de Dados Pessoais: A Função e os Limites do Consentimento. Forense, 2018.

BLASIMME, A.; VAYENA, E.; & HAFEN, E. Democratizing health research through data cooperatives. Philosophy & Technology, 31(3), pp. 473–479, 2018.

BUNDESDATENSCHUTZGESETZ [BDSG], Jan. 14, 2003, BGBL. I at 66, last amended Aug. 14, 2009, BGBL. I at 2814 (Ger.).

BUNDESVERFASSUNGSGERICHT, U. V. Census Act, BVerfGE 65, 1. 15. Dezember 1983 zum Volkszählungsgesetz 1983. Bundesanzeiger. 1983.

BRANDIMARTE, L., ACQUISTI, A., & LOEWENSTEIN, G. Misplaced confidences: Privacy and the control paradox. Social Psychological and Personality Science, 4(3), pp. 340–347, 2013.

BRASIL. Constituição (1988). Constituição da República Federativa do Brasil, 05 out. 1988. Brasília. Disponível em:

<http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm>. Acesso em: 20 jul.2019.

BRASIL. Lei Nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD), Brasília, DF, ago 2018. Disponivel em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm>. Acesso em: 05 out. 2019.

BURGIN, M. Theory of Information. Fundamentality, Diversity and Unification. World Scientific Publishing, 2010.

BYFORD, Katrin Schatz. Privacy in Cyberspace: Constructing a Model of Privacy for the Electronic Communications Environment. Rutgers Computer & Tech, 1998.

CALO, Ryan. Privacy, Vulnerability, and Affordance (February 23, 2017). Cambridge Handbook of Consumer Privacy (Evan Selinger, Jules Polonetsky, Omer Tene, eds.). Cambridge University Press: Forthcoming. Available at SSRN: https://ssrn.com/abstract=2820759

CALO, Ryan. Digital Market Manipulation. 82 Geo. Wash. L. Rev. 995, 2013.

CARBONELL, I. The Ethics of Big Data in Big Agriculture. Internet Policy Review, Vol. 5, No. 1, 2016.

CARDOSO, Renato César; GONTIJO DE OLIVEIRA, Thaís de Bessa. Conciliência e a possibilidade do neurodireito: da desconfiança à reconciliação disciplinar. Revista Brasileira de Políticas Públicas, v. 8, n. 2, pp. 946-959, 2018.

CARDOSO, Renato César. Neuroderecho y la Neurociencia del libre albedrío: una visión general. SCIO: Revista de Filosofía, [S. l.], n. 21, p. 55–81, 2021. DOI: 10.46583/scio_2021.21.843. Disponível em: https://revistas.ucv.es/scio/index.php/scio/article/view/843. Acesso em: 13 nov. 2022.

CASA CIVIL DA PRESIDÊNCIA DA REPÚBLICA, INSTITUTO DE PESQUISA ECONÔMICA E APLICADA. Avaliação de Políticas Públicas – Guia prático de análise ex ante, Volume 1. Brasília: 2018.

CASA CIVIL DA PRESIDÊNCIA DA REPÚBLICA, INSTITUTO DE PESQUISA ECONÔMICA E APLICADA. Avaliação de Políticas Públicas – Guia prático de análise ex post, Volume 2. Brasília: 2018b.

CATE, F. H. The failure of fair information practice principles. In: Winn, J. K. (Org.). Consumer protection in the age of the information economy (pp. 343–379). Ashgate, 2006. Disponível em: <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1156972>.

CLARK, C. R. Advertising restrictions and competition in the children's breakfast cereal industry. The Journal of Law and Economics 50 (4), pp. 757–780, 2007.

COCKBURN, Iain M.; HENDERSON, Rebecca; STERN, Scott. The Impact of Artificial Intelligence on Innovation: an exploratory analysis. In: Agrawal, Ajay K. et al. (Orgs.). The Economics of Artificial intelligence: an agenda. Pp. 115, 125–28, 139–43. 2019.

DONEDA, Danilo. Da privacidade à proteção de dados pessoais. Revista dos Tribunais, 2016.

ECKARD JR., E. W. Competition and the cigarette tv advertising ban. Economic Inquiry 29 (1), pp. 119–133. 1991.

ECO, U. A Theory of Semiotics, Bloomington: Indiana University Press. 1976.

EDWARDS, Lilian; VEALE, Michael, Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For. 16 Duke Law & Technology Review 18, 2017. Disponível em: <https://ssrn.com/abstract=2972855>.

EDWARDS, Lilian; VEALE, Michael. Enslaving the Algorithm: From a 'Right to an Explanation' to a 'Right to Better Decisions'?. IEEE Security & Privacy (2018) 16(3), pp. 46-54, 2018. DOI: 10.1109/MSP.2018.2701152. Disponível em: <https://ssrn.com/abstract=3052831>.

EUROPEAN DATA PROTECTION SUPERVISOR. Directive 95/46/EC. 1995.

EUROPEAN DATA PROTECTION SUPERVISOR. Annual Report. European Union, 2015. Disponível em: < https://edps.europa.eu/sites/edp/files/publication/edps_annual_report_2015_web_en.pdf>.

FISHER, William. Theories of Intellectual Property. In: S. Munzer (Org.). New Essays in the Legal and Political Theory of Property. Cambridge Univ. Press, 2001.

FLEMING, D. Can pictures be arguments? Argumentation and Advocacy, 33(1), pp. 11-22, 1996.

FLORIDI, Luciano. Introduction. In: The Online Manifesto. Being Human in a Hyperconnected Era. Floridi, Luciano (Org.). Springer, 2015.

FORUM ECONÔMICO MUNDIAL; Bain & Company, Inc. Personal data: The emergence of a new asset class. World Economic Forum, 2011.

FOUCAULT, Michel. Discipline and Punish: The Birth of the Prison. Pantheon Books, 1977.

GAL, Michal S.; AVIV, Oshrit. The competitive effects of the GDPR. Journal of Competition Law & Economics, Volume 16, Issue 3, pp. 349–391, 2020.

GALLET, C. A. The effect of the 1971 advertising ban on behavior in the cigarette industry. Managerial and Decision Economics 20 (6), pp. 299–303, 1999.

GIBSON, James J. The Ecological Approach to Visual Perception. New York: Psychology Press, 1979.

GLASER G. Barney; STRAUSS, Anselm L. The Discovery of Grounded Theory: strategies for qualitative research. New Brunswick: Aldine Transaction, 1967.

GOFFMAN, E. Relations in Public: Microstudies of the Public Order, New York: Harper Colophon, 1972.

GROARKE, L; PACLZEWISKI, C. H.; Godden, David. Navigating the visual turn in argument. Argumentation and Advocacy, v. 52, pp. 217-263, 2016.

HALLINAN, Dara; GELLERT Raphaël. The Concept of 'Information': An Invisible Problem in the GDPR. SCRIPTed Vol. 17, Issue 2, p. 269, 2020.

HARGITTAI, E. Whose space? Differences among users and non-users of social network sites. Journal of Computer-Mediated Communication, 13(1), pp. 276–297, 2007.

HARTZOG, Woodrow; RICHARDS, Neil. Privacy's Constitutional Moment and the Limits of Data Protection. Boston College Law Review, 1687, Vol. 61, Issue 5, 2020.

HAYNES, P.; CAROLYN NGUYEN, M.-H. Rebalancing socioeconomic asymmetry in a data-driven economy. In: B. Bilbao-Osorio, S. Dutta, & B. Lanvin (Orgs.). The global technology report 2014: Rewards and risks of big data, pp. 67–72, 2014.

HILDEBRANDT, M. Law as Information in the Era of Data-Driven Agency 79(1) MLR 4, 2016.

HILDEBRANDT, M. Law as Computation in the Era of Artificial Legal Intelligence. Speaking Law to the Power of Statistics (forthcoming). University of Toronto Law Journal 10, 2018.

HOLMES, V. T.; LANGFORD, J. Comprehension and recall of abstract and concrete sentences. Journal of Verbal Learning and Verbal Behavior, 15(5), pp. 559–566. 1976.

HOOFNAGLE, C. J.; URBAN, J. M. Alan Westin's privacy homo economicus. Wake Forest Law Review, 49, 261. 2014.

HOSKINS, G. T. Draft Once; Deploy Everywhere? Contextualizing Digital Law and Brazil's Marco Civil da Internet. Television & New Media, 19(5), pp. 431–447, 2017.

HUXLEY, Aldous. Brave New World. New York: Harper Brothers, 1932.

INNES, M. Signal crimes and signal disorders: notes on deviance as communicative action. British Journal of Sociology, pp. 335-55, 2004.

ISIN, Engin; RUPPERT, Evelyn. Being Digital Citizens. London: Rowman & Littlefield International, 2015.

JACOBS, Jane. The Death and Life of Great American Cities. NY: Random House, 1961.

JENSEN, C.; POTTS, C. Privacy policies as decision-making tools: An evaluation of online privacy notices. In: K. Dykstra-Erickson, M. Tsscheligi (Orgs.), Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 471–478. ACM, 2004.

JOHNSON, E. J.; BELLMAN, S.; LOHSE, G. L. Defaults, framing and privacy: Why opting in-opting out. Marketing Letters, 13(1), pp. 5–16, 2002.

JOHNSON, E. J.; GOLDSTEIN, D. G. Do defaults save lives? Science, 302(5649), 1338–1340. 2003.

JOHNSON, Garrett A.; SHRIVER, Scott K.; GOLDBERG, Samuel G. Privacy & market concentration: Intended & unintended consequences of the GDPR, 2020.

JOHNSON, Garrett A.; SHROVER, Scott K.; GOLDBERG, Samuel G. Regulating Privacy Online: An Economic Evaluation of the GDPR. Marketing Science Institute Working Paper Series 2021, Report No. 21-110. 2020b.

JØRGENSEN, Rikke Frank. Framing the Net: The Internet and Human Rights. Cheltenham: Edward Elgar, 2013.

KAFKA, Franz. Der Process. Berlin, Verlag Die Schmiede, 1925.

KAMLEITNER, B.; MITCHELL, V. W. Can consumers experience ownership for their personal data? From issues of scope and invisibility to agents handling our digital blueprints. In: J. Peck AND S. Shu (Orgs.). Psychological ownership and consumer behavior, pp. 91–118. Springer, 2018.

KJELDSEN, J. E. The study of visual and multimodal argumentation. Argumentation, 29(2), pp.115-132, 2015.

KODAPANAKKAL, Rabia I.; BRANDT, Mark J.; KOGLER, Cristoph; BEEST, Ilja van. Self-interest and data protection drive the adoption and moral acceptability of big data technologies: A conjoint analysis approach. Computers in Human Behavior, Volume 108, 2020.

KOOPS, Bert-Jaap. The trouble with European data protection law. International Data Privacy Law, Volume 4, Issue 4, pp. 250–261, 2014, https://doi.org/10.1093/idpl/ipu023.

LAKOFF, George; JOHNSON, Mark. Metaphors We Live By, 1980.

LANGER, S. K. Philosophy in a new key (3rd ed.). Cambridge, MA: Harvard University Press, 1957.

LANGSDORF, L. Image and emotion: Analyzing visual thinking. Argumentation and Advocacy, 33(1), pp. 46-52, 1996.

LESSIG, Lawrence. Reading the Constitution in Cyberspace. 45 Emory L. J. N. 3, 1996.

LESSIG, Lawrence. Code: and other laws of cyberspace. Basic Books, 1999.

LOI, Michele; DEHAYE, Paul-Olivier; HAFEN, Ernst. Towards Rawlsian 'property-owning democracy' through personal data platform cooperatives. Critical Review of International Social and Political Philosophy, 2020.

LUNA, Florencia. Elucidating the Concept of Vulnerability: Layers Not Labels, INT'L J. FEMINIST APPROACHES TO BIOETHICS, pp. 121-129, 2009.

LUPTON, Deborah. Digital sociology. London: Routledge, 2015.

LYON, David. The Electronic Eye: The Rise of the Surveillance Society. 1994.

LYON, David. Surveillance Culture: Exposure, Engagement and Ethics in Digital Modernity. International Journal of Communication 11, pp. 824–842. 2017.

MALGIERI, Gianclaudio; CUSTERS, Bart. Pricing privacy – the right to know the value of your personal data. Computer Law & Security Review: The International Journal of Technology Law and Practice, 2017, doi: 10.1016/j.clsr.2017.08.006.

MARTIN, Kirsten; NISSENBAUM, Helen. Measuring Privacy: An Empirical Test Using Context to Expose Confounding Variables, 18 COLUM. SCI. & TECH. L. REV. 176, 191. 2016.

MCGRATH, John. Loving Big Brother: surveillance culture and performance space. London: Routledge, 2004.

MCKENZIE, C. R.; LIERSCH, M. J.; FINKELSTEIN, S. R. Recommendations implicit in policy defaults. Psychological Science, 17(5), pp. 414–420. 2006.

MEDEIROS, F. A.; BYGRAVE, L. A. Brazil's Marco Civil da Internet: Does it live up to the hype? Computer Law & Security Review, 31(1), 120–130. doi:10.1016/j.clsr.2014.12.001. 2015.

MERLEAU-PONTY, Maurice. Fenomenologia da Percepção. Tradução: Carlos Alberto Ribeiro de Moura. São Paulo: Martins Fontes, 2018 (1945).

MILL, John Stuart. On Liberty. 1859.

MIRSCH, T.; LEHRER, C.; JUNG, R. Digital nudging: Altering user behavior in digital environments. In: Proceedings Der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017), pp. 634–648, 2017.

MOOJI, Joris M.; PETERS, Jonas; JANZING, Dominik; ZSCHEISCHLER, Jakob; SCHÖLKOPF, Bernhard. Distinguishing Cause from Effect Using Observational Data: Methods and Benchmarks. Journal of Machine Learning Research 17, 2016.

NEGRI, SERGIO M. C. A.; FERNANDES, E. R.; DETONI, M. R. C. Portabilidade e proteção de dados pessoais: tensões entre pessoa e mercado. Civilistica.com - Revista Eletrônica de Direito Civil, v. 1, p. 01, 2021.

NELLIS, M. (1991). The Electronic Monitoring of offenders in England and Wales: Recent Developents and Future Prospects. British Journal of Criminology, vol 31, n. 2, 1991.

NELLIS, M. Understanding the electronic monitoring of offenders in Europe: expansion, regulation and prospects. Crime, Law and Social Change, v. 62, pp. 489-510. Springer, 2014.

NIEZEN, G.; VAN DER VLIST, B. J. J.; HU, J.; FEIJIS, L. M. G. From events to goals: Supporting semantic interaction in smart environments. In Proceedings of the Computers and Communications (ISCC), 2010 IEEE Symposium on Computers and Communications, pp. 1029–1034. 2010.

OECD. Data-Driven innovation for growth and well-being: Interim synthesis report. 2014. Disponível em: <https://www.oecd.org/sti/inno/data-driven-innovation-interim-synthesis.pdf>.

OHM, Paul. Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization, 57 UCLA L. REV. 1701, 1748, 2010.

OOIJEN, I. van; VRABEC, Helena U. Does the GDPR Enhance Consumers' Control over Personal Data? An Analysis from a Behavioural Perspective. Journal of Consumer Policy volume 42, pp. 91–107, 2019.

ORWELL, George. Nineteen Eighty-Four. London: Secker and Warburg, 1949.

PALAZZO, C. Consumer campaigners read terms and conditions of their mobile phone apps. all 250,00 words. The Telegraph, 2016.

PALCZEWSKI, C. H. Argument in an off key: Playing with the productive limits of argument. In G. T. Goodnight (Org.), Communicative reason and communication communities, Vol. I, pp. 1-23. Washington, DC: National Communication Association, 2001.

PARK, Y. J. Digital literacy and privacy behavior online. Communication Research, 40(2), pp. 215–236, 2013.

PARLAMENTO EUROPEU E CONSELHO DA UNIAO EUROPEIA. Diretiva EU 2016/679 (General Data Protection Regulation – GDPR), 2016.

PRAZERES, Gustavo Cunha. Autodeterminação informacional vs. Regulação do risco: Uma abordagem sistêmica da regulamentação digital. Revista Direito e Praxis, Ahead of print, Rio de Janeiro, 2021.

PURTOVA, Nadezhda. The law of everything. Broad concept of personal data and future of EU data protection law. Law, Innovation and Technology, Volume 10, Issue 1, 2018.

REIDENBERG, J. R. Lex Informatica: The Formulation of Information Policy Rules through Technology. Texas Law Review, Vol. 76, n. 3, 1997.

REDING, V. Your data, your rights: Safeguarding your privacy in a connected world. Keynote at World Privacy Platform "The review of the EU data protection framework", 2011.

RODOTÀ, Stefano. A vida na sociedade da vigilância: a privacidade hoje. Organização, seleção e apresentação: Maria Celina Bodin de Moraes. Tradução: Danilo Doneda e Luciana Cabral Doneda. Rio de Janeiro: Renovar, 2008.

RODOTÀ, Stefano.Il mondo nella rete: Quali I diritti, quali I vincoli. iLibra Laterza, 2014.

RUBINFELD, Daniel; GAL, Michal S. Entry Barriers to Big Data, 59(2) Arizona L. Rev., 2017.

RYBCZYNSKI, Witold. Home: A Short History of an Idea. Viking Books, 1986.

SASS, T. R.; SAURMAN, D. S. Advertising restrictions and concentration: The case of malt beverages. The Review of Economics and Statistics 77 (1), pp. 66–81, 1995.

SCHWARTZ, Paul. Privacy and Democracy in Cyberspace. Vand. L. Rev. n. 294, pp. 1609-1657, 1999.

SHORE, J.; Steinman, J. Did you really agree to that? The evolution of facebook's privacy policy 2015. Resource document. Technology Science, 2015.

SLOVIC, P. Perceptions of Risk: Reflections on the Psychometric Paradigm. In: S. Krimsky; D. Goulding (Orgs.) Social Theories of Risk, Westport: Praeger, 1992.

SLOVIC, P. The Perception of Risk, New York: Earthscan, 2000.

SMITH, N. C., GOLDSTEIN, D. G.; JOHNSON, E. J. Choice without awareness: Ethical and policy implications of defaults. Journal of Public Policy & Marketing, 32(2), pp. 159–172, 2013.

SOLOVE, Daniel J. Privacy and Power: Computer Databases and Metaphors for Information Privacy. 53 STAN. L. REV. 1393, 2001.

SOLOVE, Daniel J. The Digital Person: Technology and Privacy in the Information Age. Volume 1 de Ex Machina: Law, Technology, and Society. New York: NYU Press, 2006.

SOLOVE, Daniel J.; Schwartz, Paul M. The PII Problem: Privacy and a New Concept of Personally Identifiable Information. 86 N.Y.U. L. Rev., p. 1814, 2011.

SOLOVE, Daniel J. Privacy self-management and the consent dilemma. Harvard law review, Vol. 126, pp. 1880-1903, 2013.

SOLOVE, Daniel J. The myth of the privacy paradox. GW Law School Public Law and Legal Theory, 2020. Disponível em: <https://ssrn.com/abstract=3536265>.

STAPLES, William G. The Culture of Surveillance: Discipline and Social Control in the United States. 1997.

SUNSTEIN, Cass R. Nudge. Penguin Books, 2009.

SUNSTEIN, Cass R. Human Agency and Behavioral Economics: Nudging Fast and Slow. Palgrave Macmillan Cham, 2017.

SUNSTEIN, Cass R. #Republic: Divided Democracy in the Age of Social Media. Princeton University Press, 2018.

SWEENEY, Latanya. Simple Demographics Often Identify People Uniquely 1. Carnegie Mellon Univ., Sch. of Computer Sci., Data Privacy Lab., Working Paper No. 3, 2000.

SURDEN, Harry. Machine Learning and Law. 89 Wash. L. Rev. 87, pp. 89–95, 2014.

THALER, Richard H.; SUNSTEIN, Cass R. Libertarian Paternalism. American Economic Review, vol. 93, no. 2, pp. 175-179, 2003.

TALEB, Nassim Nicholas. Anrifragile: how to live in a world we don't understand. New York: Random House, 2012.

TAYLOR, C. Modern social imaginaries. Durham: Duke University Press, 2004.

TAYLOR, Charles. A secular age. Cambridge: Harvard University Press, 2007.

TENE, Omer; POLONETSKY, Jules. Privacy in the age o big data: a time for big decisions. 64 Stan. L. Rev. Online 63, 2011.

THIEBES, Scott; LINS, Sebastian; BASTEN, Dirk. Gamifying Information Systems - a synthesis of Gamification mechanics and Dynamics. ECIS, 2014.

UNIÃO EUROPEIA. Carta dos Direitos Fundamentais da União Europeia. 2000. Disponível em: https://www.europarl.europa.eu/charter/pdf/text_pt.pdf. Acesso em: 30 maio 2022.

UNIÃO EUROPEIA. Regulamento n.º 2016/679, de 27 de abril de 2016. Regulamento Geral Sobre A Proteção de Dados Pessoais. 2016. Disponível em: https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:32016R0679&from=PT. Acesso em: 25 maio 2022.

VATANPARAST, Roxana. Designed to Serve Mankind? The Politics of the GDPR as a Global Standard and the Limits of Privacy. 80(4) Heidelberg Journal of International Law, pp. 819-845, 2020.

VILLARONGA, Eduard Fosch; KIESEBERG, Peter; LI, Tiffany. Humans forget, machines remember: Artificial intelligence and the Right to Be Forgotten. Computer Law & Security Review, Volume 34, Issue 2, pp. 304-313, 2018.

WACHTER, Sandra; MITTELSTADT, Brent; FLORIDI, Luciano. Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. International Data Privacy Law, 2017. Disponível em: <https://ssrn.com/abstract=2903469>

WACHTER, Sandra; MITTELSTADT, Brent. A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI. Columbia Business Law Review, 2019. Disponível em: <https://ssrn.com/abstract=3248829>

WEBER, Max. Economy and Society. Bedminster Press, 1978.

WEINMANN, M.; SCHNEIDER, C.; VOM BROCKE, J. Digital Nudging. Business & Information Systems Engineering, 58(6), pp. 433–436, 2016.

WHITAKER, R. The End of Privacy: How total surveillance is becoming a reality. 1999.

WRONG, Dennis H. Power: Its Forms, Bases and Uses. 1979.

ZANATTA, Rafael. Proteção de dados pessoais como regulação de risco: uma nova moldura teórica? Artigos Selecionados I Encontro da Rede de Pesquisa em Governança da Internet. Rio de Janeiro, 2017.

ZANATTA, Rafael A. F.; ABRAMOVAY, Ricardo. Dados, vícios e concorrência: repensando o jogo das economias digitais. Estud. av., São Paulo, v. 33, n. 96, pp. 421-446, 2019.

ZARSKY, Tal Z. Incompatible: The GDPR in the Age of Big Data. 47 Seton Hall L. Rev. 995, 2016.

ZUBOFF, Shoshana. Surveillance capitalism and the challenge of collective action. New Labor Forum, 2019. Disponível em: <https://journals.sagepub.com/doi/full/10.1177/1095796018819461>.

ZUBOFF, Shoshana. Caveat Usor: Surveillance Capitalism as Epistemic Inequality. In: Werbach, Kevin (Org.). After the Digital Tornado. Cambridge: Cambridge University Press, 2020. Disponível em: <https://ssrn.com/abstract=3809169>.