

A regulação estatal de aspectos econômicos da mineração de dados em *Big Data* realizada pelos provedores de aplicação e a Lei Geral de Proteção de Dados Pessoais

The state regulation of economic aspects of data mining based on Big Data carried out by the application providers and the Brazilian General Data Protection Law

Cinthia Obladen de Almendra Freitas*
Heline Sivini Ferreira**
Ricardo Cavedon***

Resumo

A mineração de dados e o *Big Data* proporcionam o tratamento de dados em contextos determinados, criando e moldando tendências em uma perene organização de complexas e dinâmicas relações sociais. A aplicação de tais técnicas faz com que o volume de dados se desvincule de sua origem e agregue características de realidade autônoma aos usuários individuais, visualizando-se nele um instrumento capaz de desequilibrar os pressupostos de uma economia de livre mercado e de gerar falhas nas relações econômicas a ponto de mitigar a livre concorrência. Por meio de método de pesquisa dedutivo e procedimento bibliográfico, apresenta-se a necessidade de estruturação e ampla proteção do meio ambiente digital com a regulação legislativa da atividade econômica de captação e tratamento de dados pelos provedores de aplicação, em especial as redes sociais, no intuito de diminuir ou eliminar as violações ao interesse social e metaindividual, tendo por base as diretrizes gerais e os princípios enunciadores da Lei Geral de Proteção dos Dados Pessoais (LGPD).

Palavras-chave: Direitos coletivos e transindividuais. Regulação estatal. Mineração de dados. Big data. Falhas de mercado. Assimetria informacional.

Abstract

Data Mining and Big Data draw a lot of attention in the rapid rise of the Internet and the digital economy, providing to organizations new tendencies at a complex and dynamic social relationship. The application of such techniques causes the volume of data to be unlinked from its origin and add autonomous reality characteristics to the individual users, visualizing in it an instrument capable of unbalancing the assumptions of a free market economy and generating relations failures mitigating the free competition. The article follows the deductive approach and the bibliographic procedure methodologies, discussing about the need for structuring and broad protection of the digital environmental with the legislative regulation of the economic activity of collecting and processing data by application providers, especially social networks, aiming at reduce or eliminate violations to interest social and meta-individual, based on the general guidelines and principles of the Brazilian General Data Protection Law.

Keywords: Collective and trans-individual rights. State regulation. Data mining. Big data. Market failures. Asymmetry of information.

1 Introdução

A sociedade pós-moderna irrompe, atualmente, com riscos e perigos inerentes e diretamente proporcionais à maior necessidade e dependência tecnológica e informacional por parte de seus indivíduos. Criam-se ambientes,

* Doutorado em Informática pela Pontifícia Universidade Católica do Paraná. Mestrado em Engenharia Elétrica e Informática Industrial pela Universidade Tecnológica Federal do Paraná. Graduação em Engenharia Civil pela Universidade Federal do Paraná. Professora Titular da Pontifícia Universidade Católica do Paraná - PUCPR para os cursos de Direito e Ciência da Computação. Professora Permanente do Programa de Pós-Graduação (Mestrado/Doutorado) em Direito. Pesquisadora do Grupo de Pesquisa - Direito do Consumo e Sociedade Tecnológica (PUCPR/CNPq). Santa Catarina – PR - Brasil. E-mail: cinthia@ppgia.pucpr.br.

** Doutora e Mestre em Direito pela Universidade Federal de Santa Catarina. Professora Adjunta no Curso de Graduação e no Programa de Pós-Graduação em Direito da PUCPR. Professora Colaboradora no curso de Pós-Graduação da Universidade Federal de Santa Catarina. Santa Catarina – PR – Brasil. E-mail: hsinivini@yahoo.com.br.

*** Doutorando do Programa de Pós-Graduação em Direito da Universidade Federal do Paraná. Mestre em Direito pela PUCPR. Pós-Graduado *Lato Sensu* em Direito Civil pela PUCPR. Especialista em Direito Aplicado pela Escola de Magistratura do Paraná. Santa Catarina – PR – Brasil. E-mail: ricardo.cavedon@gmail.com.

perspectivas e relações sociais mediante manipulações ou intencionais divulgação de dados tratados de acordo com interesses pré-determinados por provedores de aplicação de Internet. Os exemplos mais característicos são as redes sociais, cujos procedimentos de coleta e tratamento de dados criam o *Big Data*.

Mais do que um banco de dados, o *Big Data* pode ser visto como um instrumento que proporciona o tratamento de dados visando determinados contextos, criando e moldando tendências em uma perene organização de complexas e dinâmicas relações sociais. Compila e organiza tendências dos usuários, capitulando tanto questões financeiras quanto preferências pessoais e sensíveis, almejando construir uma base estruturada e organizada de perfis capaz de ser útil nas relações econômicas e mercantis. Eis a aplicação de técnicas de mineração de dados.

O *Big data* e as técnicas de mineração de dados podem ser vistos como algo a mais do que a mera compilação de dados, na medida em que seus procedimentos fazem com que o volume de dados se desvincule de sua origem (ou seja, do usuário individual), e sejam agregadas características de realidade autônoma aos usuários individuais, visualizando-se, assim, um instrumento capaz de moldar tendências e preferências a partir do ponto de vista coletivo. A sociedade como um ser coletivo experimenta influências e aspirações que quase nunca são espontâneas, mas são provocadas, incentivadas, moldadas ou, previamente, planejadas. Essa manipulação de influências coletivas, de um ponto de vista macro, é capaz de desequilibrar os pressupostos de uma economia de livre mercado e de gerar falhas nas relações econômicas a ponto de mitigar a livre concorrência. Tem-se aqui a linha mestra do artigo.

O artigo aponta para a necessidade de regulação estatal dos serviços decorrentes da mineração de dados em *Big Data* ofertados pelos provedores de aplicação, a exemplo das redes sociais. O caminho da pesquisa mostra que tal necessidade decorre de diversos aspectos: econômicos, sociais e ambientais, todos relacionados entre si. Porém, em corte metodológico, a pesquisa foi desenvolvida pelos aspectos econômicos, trazendo à discussão a existência de um desequilíbrio nos pressupostos de uma economia de livre mercado, a ponto de gerar falhas nas relações econômicas e interferir na livre concorrência.

Tal desequilíbrio é advindo das forças econômicas dos provedores de aplicação, que possuem tais forças baseadas em enorme volume de dados. Há que se ter em mente que pessoas geram dados e dados geram conhecimento sobre pessoas. E, ainda, que nunca o tratamento de dados esteve tão em pauta, seja dos pesquisadores, seja dos titulares de dados. Isto devido ao fato de que há uma relação estreita entre dados, privacidade e proteção, a ponto de se ter, no Brasil, a Lei Geral de Proteção de Dados Pessoais (LGPD), sem esquecer de outras legislações e regulamentações em nível mundial ou mesmo nacional.

Termos como coleta de dados, tratamento de dados, anonimização, bloqueio, prevenção, riscos, medidas, salvaguardas e mecanismos de mitigação de riscos, e relatórios de impacto à proteção de dados pessoais são alguns exemplos de aspectos tecnológicos que permeiam o texto legislativo da LGPD. Todos esses elementos são importantes frente ao interesse de proteger dados e garantir a privacidade como direitos fundamentais para todos os titulares de dados.

Nesse contexto, a pesquisa lança mão da privacidade como objeto central, de modo a questionar a criação de perfis (*profiling*) dos usuários, consumidores e cidadãos; apontando para os difíceis e atuais enfrentamentos, a exemplo de vazamentos de dados¹ e da existência de externalidades capazes de violar interesses coletivos quando os custos privados com o tratamento de dados não são totalmente internalizados e acabam, por consequência, sendo distribuídos e absorvidos por toda a sociedade, enfim, causando reflexos ao meio ambiente digital que deveria ser igual e equilibrado em face de todos seus destinatários.

Conseqüentemente, existem impactos na economia de mercado frente à mineração de dados em *Big Data*. Esses impactos foram explorados no artigo de modo a compreender a assimetria de informações entre os agentes econômicos, entendendo-se que assimetrias favorecem o risco moral e ação oculta, redundando na seleção adversa de tipo oculto, que acaba por desvirtuar os pressupostos do livre mercado autorregulável, sendo necessária a intervenção estatal a fim de regular as atividades econômicas e proteger a concorrência e as partes vulneráveis. No contexto trabalhado, as partes vulneráveis encontram um caminho especificamente voltado à proteção de dados pessoais e não violação da privacidade na Lei Geral de Proteção de Dados Pessoais ora em vigor no Brasil.

¹ Sobre o assunto: Vazamento em banco de dados expõe CPF de 220 milhões de pessoas (VAZAMENTO ..., 2021).

2 O meio ambiente digital e *Big Data*: a organização estratégica dos dados de usuários e os interesses da coletividade

Inicia-se contextualizando a sociedade em que se vive do ponto de vista dos dados e de sua estruturação. Conceitua-se *Big Data* de modo a refletir sobre seus reflexos em uma sociedade fundamentada no meio ambiente digital, na computação ubíqua, no paradigma *everyware* e na 4^a. Revolução Industrial.

2.1 Contextualização

A sociedade contemporânea possui características advindas da captura, indexação, identificação e segmentação de grande volume de dados existentes sobre indivíduos nas redes sociais, tratados e processados automaticamente para finalidades nem sempre divulgadas, mas quase sempre para fins comerciais, proporcionando e facilitando novas formas de vigilância, controle, discriminação e manipulação, tanto por governos quanto por empresas privadas.

Deve-se ter em mente que o volume de dados em formato digital gerado e armazenado vem crescendo exponencialmente com a evolução das tecnologias de informação e comunicação (TIC). Gantz e Reinsel (2012, p. 01) estudaram o crescimento do volume de informações no planeta, apontando que, de 2005 a 2020, o volume de dados digitais crescerá em um fator igual a 300, ou seja, passará de 130 hexabytes para 40.000 hexabytes ou 40 trilhões de gigabytes. Isto representa 5.200 gigabytes para cada homem, mulher ou criança em 2020.

Vive-se o “universo digital”, tal qual denominado por Gantz e Reinsel (2012, p. 01), os quais previram que tal universo dobrará a cada 2 anos, lembrando que esse universo compreende todo os dados digitais criados, replicados e consumidos. Melhor ainda, o universo digital é formado pelas imagens e vídeos em telefones celulares enviados ao *YouTube*, filmes digitais para TVs de alta definição, dados bancários em caixas automáticos, imagens de segurança, mensagens de voz veiculadas por linhas telefônicas digitais, e mensagens de texto (SMS ou *WhatsApp*).

Para Kuechler (2007, p. 86), aproximadamente 80% dos dados digitais encontram-se armazenados em arquivos não estruturados, sendo que parte significativa desses dados se encontra no formato de texto. “Arquivos de dados não estruturados” refere-se aos dados que, ou não têm um modelo de dados predefinido ou não estão organizados de uma maneira predefinida.

Dados estruturados ou não constituem um fator de grande interesse às organizações, possibilitando agilidade nos processos de busca e de recuperação de informações. Assim, a transformação de grandes volumes de dados textuais não estruturados em informação útil fornece elementos para a reorganização, avaliação, utilização, compartilhamento e armazenamento do conhecimento gerado a partir do conjunto bruto de dados.

As palavras de ordem da sociedade contemporânea são velocidade, mobilidade e acessibilidade. Não se pode imaginar um mundo sem a tecnologia, visto que ela está presente em todos os momentos, em todos os lugares. Fronteira, distância e informação são palavras que mudaram seus próprios significados nas últimas décadas, principalmente no século XXI. As empresas e as pessoas trabalham, vivem e se desenvolvem sem perceber, a cada minuto, que os países não possuem fronteiras no mundo digital. Pode-se adquirir um produto sentado em uma praça pública por meio de um *site* que não se sabe onde está hospedado, sendo que o produto comprado pode vir de uma terceira localidade ainda mais desconhecida. A Internet, a partir da década de 1990, vem tornando o mundo cada vez menor, mais acessível, mais veloz e sem fronteiras.

Não cabe ao usuário, da Internet e dos sistemas informatizados de um modo geral, preocupar-se com todos esses detalhes, o que interessa é realizar o objetivo em poucos cliques do *mouse* ou por meio da escolha de opções em *menus* devidamente preparados para a aplicação em questão, seja uma compra *on-line* (*e-commerce*), seja um acesso para verificar o saldo da conta corrente (*e-banking*), seja, simplesmente, trocar mensagens (*email*) ou postar fotos em redes sociais (*social networking*).

Outra palavra pode, então, ser integrada à realidade atual, qual seja: a inteligência. Vive-se em um mundo de cartões inteligentes (*smart card*), celulares inteligentes (*smartphone*) e sistemas artificiais inteligentes (*artificial intelligence* - IA). Como exemplos de sistemas inteligentes, pode-se citar: sistemas de reconhecimento de face, sistemas inteligentes de trânsito, jogos de diversão (*videogame*), sistemas de controle de qualidade, entre outros. Nesse cenário, cabe mencionar a computação ubíqua, a computação pervasiva e, ainda, a inteligência ambiental (WEISER, 1991; 1993; MOUTINHO, 2010). Chega-se, portanto, ao trinômio: *software* (programas e sistemas computacionais), *hardware* (equipamentos, dispositivos, periféricos) e *peopleware* (usuários) (LAUDON e LAUDON,

1999, p. 05-07), sendo que Greenfield (2006, p. 09) caracteriza a computação ubíqua como um único paradigma, denominado “*everyware*”, a saber:

Ainda mais invasivo, mas difícil de perceber, é que o computador saiu da mesa de trabalho e se insinuou em nossa vida diária. Essa tecnologia da informação ubíqua – ‘*everyware*’ – aparecerá em muitos contextos diferentes e tomará uma variedade de formas, mas afetará quase todos nós, estejamos ou não conscientes disto.²

Eis aqui o paradigma que norteia o desenvolvimento de um meio ambiente digital, de acordo com Cavedon *et al.* (2015), que congrega ubiquidade, pervasividade e inteligência, sendo o elemento de ligação formado pelo conjunto de dados veiculados, capturados, tratados e armazenados. Na visão de Schwab (2016), vive-se a 4.^a Revolução Industrial, uma vez que três razões sustentam tal revolução: velocidade, amplitude e profundidade, e, por último, impacto sistêmico. O autor explica que as mudanças estão ocorrendo em um ritmo exponencial e não linear, como tradicionalmente se busca descrever e entender a tecnologia. Além disso, a combinação de várias tecnologias (multiplataformas, multitarefas, entre outros) e a transformação de sistemas inteiros (desde países até empresas) “não está modificando apenas o ‘o que’ e o ‘como’ fazemos as coisas, mas também ‘quem’ somos” (SCHWAB, 2016, p. 13).

É nesse contexto de transformação que o tratamento de dados, sejam estruturados, sejam semiestruturados, sejam não estruturados, passa a assumir vital importância desde a esfera governamental até o usuário simples da Internet. Vive-se o *Big Data* (KITCHIN, 2014). O volume de dados coletado é enorme, a ponto de caracterizar o *Big Data*, possibilitando o tratamento de dados de forma quase que infinita, uma vez que se coletam inclusive dados que caracterizam o próprio conteúdo ou relacionados a ele, por exemplo, a localização de uma imagem e de quem está marcado na imagem.

A atenção recai sobre o uso de conteúdos, como divulgados pelas empresas, que podem ser “aperfeiçoado por nós” (FACEBOOK, 2016a, p. 01; 2016b, p. 01), não esclarecendo o que vem a ser exatamente esse aperfeiçoamento e que técnicas são aplicadas a partir dos dados coletados. Tecnicamente, entende-se que o texto se refere à mineração de dados, tema central neste artigo. Outro ponto controverso é que as empresas explicitam em seus Termos de Uso e Políticas de Privacidade a venda de dados aos parceiros comerciais da empresa originária dos dados.

Ao aderir a qualquer produto ou serviço *online*, é evidente a intenção do provedor do serviço ou aplicativo em capturar para si um volume de dados pessoais ou não, sendo que todos os textos, da Política de Privacidade ou dos Termos de Uso, consideram que o usuário lê e compreende a complexidade dos mesmos. Na prática, o usuário “clica” em um botão de “Aceito”, mas dificilmente lê os longos termos e políticas atentamente ou possui conhecimento técnico para desvendar os meandros da linguagem utilizada pelas empresas.

Percebe-se, portanto, que o conhecimento dos riscos e seus reflexos cada vez mais cabe ao usuário por meio dos termos e políticas disponibilizados pelos prestadores de serviços *online*. O hábito de não ler e simplesmente “aceitar” os termos e políticas precisa ser modificado diante da divulgação dos riscos e da discussão do conteúdo dos documentos ofertados pelos provedores aos usuários. Cabe ao usuário ter consciência e responsabilidade pelo que ele aceita e, assim, entender a complexa coleta de dados pessoais ou não e, principalmente, entender que os provedores irão, com certeza, processar e minerar dados e descobrir informações automaticamente ou derivadas. Porém tais informações podem e irão “contar” ou “dizer” muito a respeito de cada usuário.

A preocupação recai sobre a utilização indiscriminada do *Big Data* para fins lucrativos, a qual pode gerar ineficiência no livre desenvolvimento de mercado a partir da assimetria de informações entre os agentes econômicos, haja vista outorgar uma posição relevante e dominante para o seu detentor, permitindo que exerça de forma abusiva tal posição, ocasionando uma infração à ordem econômica.³

² Texto original: “*Ever more pervasive, ever harder to perceive, computing has leapt off desktop and insinuated itself into everyday life. Such ubiquitous information technology – ‘everyware’ – will appear in many different contexts and take a wide variety of forms, but it will affect almost every one of us, whether we’re aware of it or not.*”

³ A ordem econômica é um bem jurídico de característica transindividual, protegida, atualmente, tanto pela Lei da Ação Civil Pública (art. 1º, inciso V, da Lei n.º 7.347/85), como pela regulação legal disposta na Lei n.º 12.529, de 30 de novembro de 2011, que, em seu artigo 36, elenca as infrações a este bem jurídico transindividual, assim dispendo: Art. 36. Constituem infração da ordem econômica, independentemente de culpa, os atos sob qualquer forma manifestados, que tenham por objeto ou possam produzir os seguintes efeitos, ainda que não sejam alcançados: I - **limitar, falsear ou de qualquer forma prejudicar a livre concorrência** ou a livre iniciativa; II - dominar mercado relevante de bens ou serviços; III - aumentar arbitrariamente os lucros; e IV - **exercer de forma abusiva posição dominante**. Cabe ao CADE - Conselho Administrativo de Defesa Econômica – na esfera administrativa decidir sobre a existência de infração à ordem econômica e aplicar as penalidades previstas em lei (art. 9º, inciso II, da Lei n.º 12.529/2011).

2.2 *Big Data*: conceito e reflexos na sociedade informacional

Os dados são plurais e provêm desde uma lombada eletrônica até a geladeira, passando pelos *emails* e páginas *web*. Portanto, os sistemas computacionais estão sendo mais exigidos para tratar, armazenar, indexar e recuperar dados, visto que os sistemas tradicionais de processamento de dados tendem a colapsar. Entra em foco o *Big Data*.

O atual tratamento de dados somente é possível frente ao acúmulo de dados em proporções antes inimagináveis. O conceito que descreve esse volume de dados vem sendo conhecido como *Big Data*:

Os dados das redes sociais online podem ser usados para extrair informações sobre padrões de interações interpessoais e opiniões. Esses dados podem auxiliar no entendimento de fenômenos, na previsão de um evento ou na tomada de decisões. Com a ampla adoção dessas redes, esses dados aumentaram em volume, variedade e precisam de processamento rápido, exigindo, por esse motivo, que novas abordagens no tratamento sejam empregadas. Aos dados que possuem tais características (volume, variedade e necessidade de velocidade em seu tratamento), chamamo-los de *big data* (FRANÇA *et al.*, 2014, p. 8).

Pelo conceito apresentado, entende-se que praticamente todo conteúdo produzido no âmbito das redes sociais e dos aplicativos de *smartphones* e *tablets* pode ser considerado *Big Data*, o qual possui três V's, de acordo com Kitchin (2014, p. 68), como características intrínsecas, a saber: volume, velocidade e variedade de tipos de dados, estruturados ou não, sendo frequentemente referenciados temporal e espacialmente.

Big Data tem sido considerado muito mais do que grandes quantidades de dados produzidas por *smartphones* e redes sociais. Manyika *et al.* (2011, p. vi) apontam que o *Big Data* alcança qualquer setor ou função da economia global: "*Big data - large pools of data that can be captured, communicated, aggregated, stored, and analyzed - is now part of every sector and function of the global economy.*". Os autores apresentam que:

Each second of high-definition video, for example, generates more than 2,000 times as many bytes as required to store a single page of text. In a digitized world, consumers going about their day - communicating, browsing, buying, sharing, searching - create their own enormous trails of data (MANYIKA *et al.*, 2011, p. 01).

E, portanto, os autores definem o *Big Data* de maneira subjetiva, mas ampla o suficiente para realmente englobar o que é *Big Data*: "*refers to datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze*". Não há limitação de tamanho, tempo e lugar para *Big Data*. A visão dos autores é positiva, no sentido de que existem evidências de que *Big Data* pode desempenhar um papel econômico significativo para o benefício, não somente do comércio privado, mas também das economias nacionais e seus cidadãos (MANYIKA *et al.*, 2011, p. 01).

Os dados podem agregar valor substancialmente à economia mundial, sendo que Manyika *et al.* (2011, p. 02) demonstram que o valor dos dados está em sua aplicação criativa e eficaz nos mais variados setores. Por exemplo, ao considerarem os cuidados de saúde nos Estados Unidos, os autores estimam que o valor potencial dos dados no setor poderia alcançar mais de US\$ 300 bilhões a cada ano. Já no setor privado, os autores estimam que um varejista que explore *Big Data* tem o potencial de aumentar sua margem operacional em mais de 60%. Em relação à Europa, os autores estimam que a administração governamental pode economizar mais de € 100 bilhões em melhorias de eficiência operacional (redução de fraudes, erros e disparidades fiscais) se for aplicado *Big Data*. Assim, eles demonstram que (MANYIKA *et al.*, 2011, p. 02):

There are many ways that big data can be used to create value across sectors of the global economy. Indeed, our research suggests that we are on the cusp of a tremendous wave of innovation, productivity, and growth, as well as new modes of competition and value capture - all driven by big data as consumers, companies, and economic sectors exploit its potential.

As palavras de ordem são: inovação, produtividade e crescimento. Sem deixar de mencionar a convergência. A computação ubíqua, a computação pervasiva e a computação inteligente alcançaram o ponto de desenvolvimento em que podem ser trabalhadas de maneira integrada, associada e colaborativa. Afirma-se que, para trabalhar com *Big Data* e tecnologias convergentes, os Estados Unidos da América possuem uma escassez "de 140.000 a 190.000 pessoas com profundas habilidades analíticas, bem como 1,5 milhão de gestores e analistas para

analisar *Big Data* e tomar decisões com base em suas descobertas” (MANYIKA *et al.*, 2011, p. 03). Uma vantagem ainda é a real possibilidade de se aplicar análises sofisticadas que podem melhorar substancialmente a tomada de decisões, minimizar riscos e descobrir conhecimentos valiosos que, de outra forma, permaneceriam ocultos (MANYIKA *et al.*, 2011, p. 05).

Mas existem problemas a ponto de Manyika *et al.* (2011, p. 15) questionarem: “*Is big data simply a sign of how intrusive society has become, or can big data, in fact, play a useful role in economic terms that can benefit all societal stakeholders?*”, mencionando a sociedade intrusiva que surgiu com o alto grau de vigilância (*surveillance*), que não somente atinge cada indivíduo, mas tornou-se uma faceta da vida em um mundo conectado e globalizado (MARX, 2004. p. 18-19).

Esta é a realidade: dados e mais dados. Vive-se imerso em dados que são coletados, capturados, processados e analisados. Nesse cenário, *Data-Pop Alliance* (BHARGAVA *et al.*, 2015, p. ii) apresenta que *Big Data* define um ecossistema resultante do surgimento concomitante de três C’s:

- a) *Digital Crumbs* – partes de dados passivamente emitidos e/ou coletados por dispositivos digitais que constituem conjuntos e fluxos de dados de grande volume e contêm informações únicas sobre comportamentos e crenças;
- b) *Big Data Capacities* - conjunto de ferramentas e métodos, hardware e *software*, *know-how* e habilidades, necessários para processar e analisar esses novos tipos de dados - incluindo técnicas de mineração de dados, descoberta do conhecimento, visualização, aprendizagem e algoritmos estatísticos, entre outros;
- c) *Big Data Communities* - descrevem os vários atores envolvidos no ecossistema *Big Data*, desde geradores de dados até seus analistas e usuários finais, ou seja, potencialmente toda a população no sentido estatístico.

Por consequência de todas essas características e vantagens, o *Big Data* tem sido comparado ao ‘novo óleo da economia digital’, sendo o nível de transformação e refinamento dos dados tão relevante que o fenômeno *Big Data* tem sido definido como “uma mentalidade (...) para transformar a bagunça em sentido” (BHARGAVA *et al.*, 2015, p. 05). Entende-se, portanto, que *Big Data* é muito mais do que um grande volume de dados. *Big Data* inclui o poder de tratamento e análise de dados semiestruturados e não estruturados provenientes de diversas fontes, formatos, estruturas e tipos. Consequentemente, este permite revelar tanto as formalidades quanto as restrições dos bancos de dados convencionais, de modo a ampliar as análises interativas e exploratórias sobre quaisquer conjunto de dados capturados e armazenados. O importante é que o *Big Data* tem características que permitem preservar a fidedignidade dos dados, ou seja, a autenticidade dos dados. Cabe, assim, resumir os três V’s apresentados por Kitchin (2014, p. 68), visto que fundamentam as características intrínsecas do *Big Data*, a saber: volume, velocidade e variedade de tipos de dados.

Sobre o primeiro ‘V’, volume de dados, Schneider (2016, p. 01) afirma que, em sendo os dados criados na atualidade em sua maioria dados não estruturados, o não entender tais dados fará com que as empresas venham a perder inteligência digital. Schneider (2016, p. 01) aponta que aproximadamente 2,5 bilhões de Gigabytes de dados são criados por dia. É aqui que se adentra ao segundo ‘V’, a velocidade de criação de dados, podendo-se mencionar a criação em tempo real (*real time*) dos dados frente à quantidade de sensores e câmeras associados às redes sociais e aplicativos *online*, fazendo com que todos esses elementos trabalhem concomitantemente e ininterruptamente. O conceito de tempo proporciona ao *Big Data* a característica de dinamicidade à natureza da criação dos dados. Além disto, o volume e a velocidade decorrem das repetições de observações realizadas no tempo e no espaço (KITCHIN, 2014, p. 76).

Já o terceiro ‘V’, variedade de tipos de dados, é que fornece ao *Big Data* a capacidade de não somente armazenar dados, mas de processá-los. Cabe ressaltar que a classificação de dados é extensa e envolve conceitos teóricos e técnicos, antes mesmo de conceitos e aplicações jurídicas do termo ‘dados’. Como referencial teórico para a discussão sobre classes de dados recomenda-se tanto material bibliográfico das áreas de Informática e Estatística quanto da área técnica de *Big Data*. O termo ‘dados’ é tão amplo que permite até mesmo ser utilizado conceitualmente como política e fenômeno social, tal qual mencionado por *Data-Pop Alliance* (BHARGAVA *et al.*, 2015, p. ii): “*data is changing the world*”, a ponto de se considerar o conceito de ecossistema de dados (*data ecosystems*), envolvendo organizações complexas de relações sociais dinâmicas por meio das quais dados e informações se movem e se transformam.

Kitchin (2014, p. 1-66) mostra que os dados podem se apresentar de diferentes formas (quantitativos e qualitativos), estruturas (estruturados, não estruturados e semiestruturados), fontes (capturados, derivados, *exhaust*, transitório), origens (primário, secundário, terciário) e tipos (indexais, atributos, metadados). Além disto, dados podem se apresentar, em termos de materialidade, por meio de diferentes formatos, ou seja, números, textos, símbolos, imagens, sons e ondas eletromagnéticas. Deve-se lembrar que um espaço em branco e a ausência de som ou ruídos, o silêncio, também são formatos de dados (KITCHIN, 2014, p. 4). Há que se incluir na vasta imensidão dos dados o conceito de dados abertos (*open data*) (OPEN DATA HANDBOOK, 2017, p. 1). Portanto, na sociedade contemporânea e informacional, a variedade de dispositivos gera pluralidade de dados, os quais não mais se encontram organizados e estruturados pelos modelos convencionalmente conhecidos.

3 Os impactos do *Big Data* na economia de mercado

Por se tratar de uma complexa organização de dados, que vai muito além do mero armazenamento de dados, mas alcança a interpretação e a depuração de complexa dinâmica da fenomenologia social, o *Big Data* desvincula-se de suas origens individualizáveis e alcança a noção de uma realidade autônoma capaz de produzir reflexos em direitos da coletividade e agregar à dimensão social um aspecto transindividual inerente à sociedade contemporânea.

Na medida em que sua utilização gera desequilíbrios na economia de livre mercado, proporcionando falhas a exemplo da assimetria informacional entre seus destinatários, ressurge daí a compreensão de que viola interesses coletivos e transindividuais.

Os direitos coletivos e transindividuais detêm estrita correlação com as falhas de uma economia de mercado autorregulado. Partindo da premissa de que os direitos individuais foram instrumentalizados e instituídos para fomentar a existência de uma economia autorregulável, para direcionar a economia aos pressupostos capitalistas, pode-se igualmente definir que as falhas nesse mercado se materializam em mitigações a direitos coletivos, vistos em sua titularidade indeterminada e em seu objeto indivisível, ou seja, violam interesses da sociedade como um todo, inclusive das futuras gerações. Eis o confronto entre a tecnologia e os direitos fundamentais.

Associadas à perda do bem-estar coletivo, as falhas econômicas podem ser identificadas na formação de concentrações de mercado, que geram ineficiências e desequilibram o livre desenvolvimento da livre concorrência. A ineficiência estática ocorre porque a concentração de mercado traduziria altos preços (nem sempre para o destinatário imediato) em contraponto a um baixo custo de exploração, derivado da pouca oferta ou da quase nula concorrência que promoveriam, que redundaria em lucros exacerbados a um único produtor e concentrador capitalista, no caso, as empresas provedoras de aplicação. Em contraponto, dado a facilidade de manutenção pelo único fornecedor monopolista da grande demanda pelos seus produtos, a ineficiência dinâmica tornaria quase nulo o incentivo ao desenvolvedor para investir em novas tecnologias e inovações técnicas, dado que apenas a livre concorrência faz com que o produtor inove e invista para buscar lucros e outros mercados consumidores (CAVEDON, 2015, p. 162).

A exploração da livre concorrência no mercado pressupõe agentes econômicos dotados de atributos equidistantes. A livre concorrência é um bem jurídico protegido constitucionalmente, nos termos do art. 170, inciso IV, da Constituição Federal, e constitui, com a liberdade de iniciativa, um dos mais importantes pilares da ordem econômica em que se desenvolve o Estado democrático brasileiro (GRAU, 2008). Portanto, quando um agente econômico com sua atividade empresarial opera no mercado e lança consequências em destinatários indeterminados, como no caso do uso indiscriminado para fins comerciais do *Big Data*, ocorre a automática violação da livre concorrência e, por consequência, a direitos coletivos de características transindividuais.

A livre concorrência de mercado pressupõe que haja uma simetria entre as informações de e para os agentes econômicos, notadamente acerca das características e tendências dos destinatários econômicos. Na medida em que a assimetria de informação pode gerar o que se denomina de risco moral e ação oculta, redundando na seleção adversa de tipo oculto, que acaba por desvirtuar os pressupostos do livre mercado autorregulável, é necessária a intervenção estatal a fim de regular as atividades econômicas e proteger a concorrência e as partes vulneráveis.

Ballart e Matas (2002, p. 482-487) elencam situações quando o modelo básico econômico, que, na maioria dos países capitalistas, é o de livre mercado, afasta-se do pressuposto para seu livre desenvolvimento, mitigando a competitividade. Quando isto ocorre, aumenta-se concomitantemente a dificuldade da permanência da eficiência de Pareto, gerando mitigação do bem-estar coletivo. Tipificam esses autores as falhas de mercado em “externalidades

negativas, monopólios naturais e assimetrias na informação”, além da “provisão de bens públicos” (MICHELL; SIMMONS, 2003, p. 33.68; PEREIRA, 2005).

As características do *Big Data*, quando utilizado indiscriminadamente para fins comerciais, bem como os seus reflexos na sociedade informacional, contribuem para a assimetria de informações, fragilizando o mercado e expondo os usuários à formação de um espaço que se desvincula dos usuários individuais a partir da sistematização e organização estratégica de seus dados pessoais. No *Big Data* não existem indivíduos, mas sim um volume de dados desvinculado de sua origem, processável e comercializável pelas empresas, mas não de seu exclusivo pertencimento, na medida em que seu uso interessa também a toda a coletividade, inclusive às futuras gerações.

Quando do uso indiscriminado desse volume de dados estrategicamente trabalhadas para finalidades nem sempre visando apenas o bem-estar coletivo, mas apenas o lucro, viola-se um direito coletivo de características transindividuais. É latente a necessidade de razoável utilização desses dados não pertencentes exclusivamente à empresa que os detém, mas a toda a sociedade, inclusive àqueles que futuramente poderão vir a ter contato ou ser expostos aos reflexos do uso indiscriminado dos dados. Passa-se, nessa perspectiva, a visualizar o *Big Data* como uma realidade autônoma, desvinculada dos dados individuais que lhe deram origem e cujos reflexos atingem a sociedade como um todo, inclusive, em maior ou menor medida, as futuras gerações.

Disto extrai-se o questionamento: O uso exclusivo e desarrazoado desse conjunto de dados, previamente tratados e economicamente relevantes, viola ou não de modo visceral o interesse coletivo capaz de mitigar os pressupostos de uma economia autorregulável, gerando desequilíbrios na fruição e desenvolvimento do livre mercado existente na maioria dos países capitalistas? A partir do momento que o *Big Data* é formado, desvinculando-se estrategicamente o conjunto de dados de seus usuários individuais, o seu correto uso e exploração, ainda que para fins econômicos, materializa-se como realizador de um direito coletivo, de característica transindividual, porquanto seu adequado direcionamento interessa a uma titularidade indeterminada de destinatários, não obstante seu suporte material continue a pertencer tão somente à empresa que coletou os dados. Tal desvinculação é possível por exigências das técnicas de mineração de dados e por técnicas específicas de anonimização.

Essa gama de dados não pode mais ser vista individualmente, mas como um todo suscetível a determinadas finalidades, na maioria das vezes, sendo usada exclusivamente no intuito comercial e lucrativo, gerando externalidades negativas e consequências muitas vezes nefastas para aqueles usuários individualizados que detém uma porção indivisível de interesse na sua correta aplicação. Os dados não mais podem ser divididos e individualizados, mas apenas usados coletivamente e, muitas vezes, comercializados, como a imensa base de dados estrategicamente organizada, anonimizada e tratada. Tal asserção é relevante, pois é necessário se pensar e regular o uso das novas tecnologias e desse conjunto indivisível de dados, de modo a não violar o interesse social existente em sua adequada destinação.

Assim, o uso do *Big Data* para interesses econômicos daqueles que detém seu controle amplia a assimetria informacional entre empresas e usuários. Tal circunstância relativiza as posições anteriormente equidistantes no mercado de consumo, mitigando a liberdade de seus agentes e, em última instância, a própria livre concorrência. Na medida em que um dos agentes isoladamente possui mais informações sobre o outro polo da relação econômica, mais se utiliza dessas informações não só para conhecer as características e preferências dos agentes econômicos (consumidores, por exemplo), mas, por vezes, para, em um processo reverso, “moldar, modificar ou mesmo manipular as suas opiniões e crenças” (AB2L, 2018).

Assim, verificando-se os reflexos dessas falhas de mercado, em que se mitigam um dos atributos necessários para que o mercado se desenvolva legitimamente, em última instância violando a livre concorrência, é possível se aferir a mitigação ou violação de um interesse coletivo de característica transindividual. E, lembrando, a sociedade anseia por ter sempre o bem-estar coletivo acima de interesses mercadológicos de determinados agentes econômicos, assim, a correção dessas falhas de mercado se fazem mediante a regulação estatal, por meio de aparato legislativo, de modo que as externalidades desse comportamento mitigam, inegavelmente, o bem-estar coletivo e, em última análise, violam direitos da coletividade no seu aspecto metaindividual.

Associadas à perda do bem-estar coletivo, as falhas da economia de livre mercado passam a ser identificadas como mitigadoras do interesse coletivo ou violadoras de direitos transindividuais. São verdadeiros elementos iniciadores, de onde se extrai a justificativa estruturante para a caracterização dos direitos coletivos (CAVEDON, 2015). É possível, assim, dizer que todos os bens jurídicos, quando envolvem tutela de direitos coletivos (e não

mera tutela coletiva de direitos), estão em maior ou menor grau associados a uma dessas falhas de mercado, geradoras de ineficiências e diminuição do bem-estar coletivo.

Portanto, se no mercado existir um agente econômico que, em posição privilegiada dos demais, domina informações sobremaneira diferenciadas, advindas da organização estratégica de dados capturados a partir de usuários de serviços *online*, a exemplo das redes sociais e demais aplicativos, mesmo que gozem de autorização e consentimento dos usuários individuais por meio de contratos de adesão aos Termos de Uso e Políticas de Privacidade, o uso indiscriminado desses dados para fins comerciais sem controle ou regulação estatal gera a chamada falha de mercado denominada de assimetria da informação.

Outro fator que gera violação do interesse coletivo pode ser explicado pelo conceito econômico de externalidade negativas oriundas da atividade econômica. No que pertine à utilização do *Big Data* pelos agentes econômicos, é possível se ponderar que a sua indiscriminada fatorização para segmentar os usuários, caracterizando perfis (*profiling*) (FERRARIS *et al.*, 2013), e, por consequência, segmentar o mercado de consumo, redundando em externalidades que mitigam o bem jurídico transindividual do meio ambiente digital (CAVEDON *et al.*, 2015).

As externalidades são capazes de violar esses interesses coletivos quando os custos privados com o tratamento de dados não são totalmente internalizados e acabam, por consequência, sendo distribuídos e absorvidos por toda a sociedade, gerando violações à intimidade e à privacidade dos usuários, com discriminação e caracterização de perfil social (*profiling*) de destinatários (usuários e consumidores), segmentando o meio ambiente digital, que deveria ser igual e equilibrado em face de todos seus destinatários.

Outra externalidade é a deterioração dos demais fatores de produção e dos outros agentes econômicos, na medida em que o *profiling* redundando, em última análise, no prévio direcionamento de demandas consumeristas para grupos, segmentos ou camadas sociais, discriminando toda uma gama de agentes econômicos ante o processo produtivo desencadeado de acordo com os dados cujo controle é somente de um único ator em posição dominante. O poder então conferido às empresas que se beneficiam dos dados de seus usuários, a ponto de formarem *Big Data*, é demasiado suficiente ao controle de processos coletivos e necessidades sociais.

Há nesse contexto a perda do bem-estar coletivo, o que gera uma fissura à ideia da mão invisível do mercado, pregada por Adam Smith (SMITH, 2014), no dogma clássico de que todos contribuem para o bem-estar geral quando pensam estar trabalhando para seus interesses próprios. Aqui os agentes econômicos ao buscarem seus interesses próprios acabam por gerar involuntariamente violações aos interesses transindividuais ou coletivos, cabendo ao ente estatal regular o serviço de interesse coletivo.

Tal contexto alcança não só os agentes econômicos envolvidos, mas toda a população e, inclusive, as futuras gerações. Pense-se na formação de bolhas informacionais (PARISER, 2012), na fragmentação das informações oriundas do *Big Data*, a ponto de se identificar previamente cada usuário, tornando-o cada qual involuntariamente parte de um perfil (*profiling*) previamente definido, podendo-se até mesmo estabelecer parâmetros discriminatórios, os quais alcançam a violação não somente de aspectos transindividuais e sociais, mas também a violação da personalidade individual da liberdade de cada usuário.

Essas razões, aliadas à necessidade de estruturação e proteção ampla do meio ambiente digital, são os motivos que levam à necessidade de regulação da atividade econômica do controle de dados capturados pelos serviços *online*, a exemplo das redes sociais, no intuito de diminuir ou eliminar essas falhas no desenvolvimento da economia, na tentativa de maximizar o bem-estar social coletivo e proteger os interesses metaindividuais, em regra indivisíveis e direcionadas a destinatários indeterminados.

4 Necessidade de regulação estatal do *Big Data* e a Lei Geral de Proteção de Dados Pessoais

A necessidade de se construir juridicamente o aspecto transindividual do *Big Data* como advinda da normatividade de um bem jurídico coletivo, a exemplo do meio ambiente digital (CAVEDON *et al.*, 2015), e então se instituir uma limitação à livre disponibilidade privada dos dados pertencentes às empresas de TIC, é latente, e de extrema relevância na atualidade. Não se trata somente de uma base de dados, mas vai além, visto envolver um conjunto organizado e sistematicamente aplicado de técnicas que fazem seu portador deter relevante vantagem na alocação de interesses. *Big Data* vai além do conceito de um banco de dados tradicional, sendo realidade autônoma, haja vista que é estrategicamente organizado e tratado, ostentando uma sistematização ímpar capaz de alcançar *status* normativo autônomo.

Quando se confere *status* normativo autônomo a um direito transindividual, como com os reflexos do *Big Data*, em regra visto como um bem incorpóreo, de plano reconhece-se automaticamente a limitação à atividade econômica que lhe serve como suporte material, no caso restrição ou regulação à utilização dos dados individuais organizados e economicamente relevantes, mas indissociáveis em seu conjunto. Nesse caso, reconhecendo-se a normatividade da proteção do equilíbrio do meio ambiente digital, admite-se de forma automática a limitação e regulação da atividade econômica das empresas de TIC, feita, nesse caso, por meio de lei.

Desse modo, tem-se que a liberdade de tratamento e exploração de dados individuais capturados pelas empresas de TIC, a exemplo das redes sociais, pautada pela livre promoção e desenvolvimento de interesses individuais e comerciais, não lograria por si só nem mesmo reflexamente conter a violação a bens jurídicos metaindividuais, como o meio ambiente digital, notadamente diante das externalidades negativas e de assimetria de informações verificadas pelo uso indiscriminado e estratégico (para fins comerciais) da variedade do objeto primordial do *Big Data*, ou seja, os dados. Reforça-se, então, a necessidade de regulação e fiscalização dessa atividade econômica, nos termos dos artigos 170 e 174 da Constituição da República.⁴

Segundo a AB2L (2018), a questão da regulação da tecnologia é cada vez mais complexa no mundo atual, notadamente porque os bens mais importantes tendem a se tornar de características imateriais, sendo as titularidades individuais paulatinamente substituídas por acessos compartilhado, daí se tornando multifuncionais e interconectados, havendo interações cada vez mais complexas, intensas e abrangentes.

O entendimento técnico do *Big Data* e de seus reflexos jurídicos na sociedade informacional assume uma preocupação cada vez mais latente. A característica que assume de um direito coletivo, de natureza transindividual, ostenta titularidade indeterminada e objeto indivisível, unidos por uma relação jurídica base com a parte contrária (a empresa), tendendo a levar à conclusão da necessidade de regulação, o que invariavelmente restringe direitos individuais, seja das empresas, seja de indivíduos. Assim, a utilização do *Big Data* deve ser regulada pelo aparato estatal, uma vez que mitiga os pressupostos da livre concorrência.

É necessário que haja a promoção de uma legislação específica a regular e fiscalizar a livre disponibilidade dos dados pessoais dos usuários, não podendo esses dados serem estrategicamente utilizados, sem a necessária regulação, para fins de exploração comercial e caracterização de perfil (*profiling*),⁵ uma vez que os riscos envolvidos nessa atividade são absorvidos pela sociedade como um todo, podendo gerar consequências nefastas tanto para as presentes como para as futuras gerações.⁶

Em 14 agosto de 2018 foi sancionada a Lei Geral de Proteção de Dados Pessoais (Lei n.º 13.709, com redação atualizada em Lei n.º 13.853, de 08 de julho de 2019) (BRASIL, 2018) a qual

[...] sobre o tratamento de dados pessoais, inclusive nos meios digitais, por pessoa natural ou por pessoa jurídica de direito público ou privado, com o objetivo de proteger os direitos fundamentais de liberdade e de privacidade e o livre desenvolvimento da personalidade da pessoa natural (art. 1º, LGPD) (BRASIL, 2018).

Definindo tratamento de dados como (art. 5º, inciso X, LGPD):

[...] toda operação realizada com dados pessoais, como as que se referem a coleta, produção, recepção, classificação, utilização, acesso, reprodução, transmissão, distribuição, processamento, arquivamento, armazenamento, eliminação, avaliação ou controle da informação, modificação, comunicação, transferência, difusão ou extração (BRASIL, 2018).

Cabe explicar que a LGPD não é a única legislação do ordenamento brasileiro que trata de dados pessoais, mas pode ser considerada uma evolução no caminho pela proteção de dados pessoais, desde a Constituição da

⁴ Art. 170. A ordem econômica, fundada na valorização do trabalho humano e na livre iniciativa, tem por fim assegurar a todos existência digna, conforme os ditames da justiça social, observados os seguintes princípios:

(...) IV - livre concorrência; V - defesa do consumidor; (...) VII - redução das desigualdades regionais e sociais; (...) Parágrafo único. É assegurado a todos o livre exercício de qualquer atividade econômica, independentemente de autorização de órgãos públicos, salvo nos casos previstos em lei.

Art. 174. Como agente normativo e regulador da atividade econômica, o Estado exercerá, na forma da lei, as funções de fiscalização, incentivo e planejamento, sendo este determinante para o setor público e indicativo para o setor privado.

⁵ É de assustar: Facebook guarda dados pessoais que você nem lembrava... Disponível em: <https://tecnologia.uol.com.br/noticias/redacao/2017/10/16/e-de-assustar-facebook-guarda-dados-pessoais-que-voce-nem-lembrava.htm>. Acesso em: 18 nov. 2017.

⁶ Vide situações em que empresas de redes sociais e serviços *online* promovem tratamento de dados, como a caracterização de perfil (*profiling*), no intuito de explorar fraquezas humanas, tendências viciosas e capazes de serem consideradas como doenças, como por exemplo, os vícios humanos em jogos de azar ou consumo compulsivo.

República, pela qual são “invioláveis a intimidade, a vida privada, a honra e a imagem das pessoas, assegurado o direito a indenização pelo dano material ou moral decorrente de sua violação” (art. 5º, inciso X) (BRASIL, 1988), até o Marco Civil da Internet (BRASIL, 2014).

Desde 2014, o Marco Civil da Internet (Lei n.º 12.965) já continha no rol de princípios (art. 3º) o princípio da proteção da privacidade e dos dados pessoais, assegurando-o como direitos e garantias dos usuários de Internet (art. 7º) a inviolabilidade e sigilo do fluxo de suas comunicações e inviolabilidade e sigilo de suas comunicações privadas armazenadas, salvo por ordem judicial. Além disso, na Seção II, a qual trata especificamente sobre proteção aos registros, aos dados pessoais e às comunicações privadas, está previsto que a guarda e disponibilização de tais elementos “devem atender à preservação da intimidade, da vida privada, da honra e da imagem das partes direta ou indiretamente envolvidas” e, ainda, que o fornecimento de dados privados poderá ser realizado somente por ordem judicial.

Mesmo sabendo-se que as leis requerem amadurecimento contínuo, constatar as alterações e evoluções permite não somente ao legislador, mas a todos os envolvidos, compreender que a LGPD tem forte inspiração no *General Data Protection Regulation* (GDPR), vigente na União Europeia desde 2018 (UNIÃO EUROPEIA, 2016). Por conseguinte, apesar de terem suas particularidades, ambas compartilham o mesmo modelo, lógica e estrutura de um regulamento geral de proteção de dados.

A LGPD não aborda o conjunto de dados como algo indissociável e portador de uma realidade autônoma, ou seja, não aborda especificamente *Big Data*, tal qual analisado neste artigo, mas dispõe sobre a regulação do tratamento de dados com o objetivo de proteger os direitos fundamentais de liberdade e de privacidade e o livre desenvolvimento da personalidade da pessoa natural. Vislumbra unicamente os reflexos individuais e subjetivos do problema ora apresentado, olvidando-se, contudo, do aspecto objetivo e transindividual, notadamente quando em cena o equilíbrio do meio ambiente digital.

A regulação, segundo o texto da LGPD, abrange qualquer operação de tratamento realizada por meio total ou parcialmente automatizado, por pessoa natural ou por pessoa jurídica de direito público ou privado, independentemente do país de sua sede ou do país onde estejam localizados os dados, desde que a operação de tratamento seja realizada no território nacional (art. 3º, inciso I, LGPD) (BRASIL, 2018); ou os dados pessoais objeto do tratamento tenham sido coletados no território nacional (art. 3º, inciso III, LGPD) (BRASIL, 2018). Corretamente, sob os aspectos técnico e jurídico, a LGPD elenca que basta os dados terem sido coletados no território nacional para incidir a legislação, o que não obstante pode ser suscetível de diversas interpretações com intenção de fugir à sua incidência. Ela exclui de sua abrangência os tratamentos de dados realizados por pessoa natural para fins exclusivamente particulares e não econômicos, ou seja, para fins exclusivamente pessoais; ou realizados para fins exclusivamente jornalístico, artístico ou acadêmico (art. 4º, LGPD) (BRASIL, 2018).

A LGPD também elenca o uso compartilhado de dados (art 5º, inciso XVI) (BRASIL, 2018), sendo que qualquer atividade de tratamento de dados deve estar pautada na boa-fé e em princípios (art. 6º) (BRASIL, 2018), tais como: finalidade, adequação, necessidade, livre acesso, qualidade dos dados, transparência, segurança, prevenção, não discriminação e responsabilização e prestação de contas. Todos pautados pelo devido informar ao titular, sempre considerando o contexto de tratamento, o qual deve se limitar ao mínimo necessário para a realização das finalidades almejadas, de modo a permitir que o titular tenha consulta garantida, gratuita e facilitada sobre as modalidades de tratamento e sobre a integralidade dos seus dados pessoais. Há que se preocupar em garantir deste a exatidão, a clareza e a atualização dos dados até a segurança da informação, área que passará a ser de extrema importância para qualquer pessoa natural ou jurídica de direito público ou privado que realize tratamento de dados pessoais. Para tal, o *compliance* poderá potencializar a efetividade da lei, frente aos princípios de transparência, prevenção e responsabilização e prestação de contas.

Importante destacar, entre os princípios, a não discriminação, que terá papel fundamental em regular o *Big Data* enquanto realidade autônoma, para evitar e limitar a atividade econômica, a ponto de colocar em risco direitos transindividuais, de acesso igualitário aos mesmos ambientes digitais e à mesma exposição às informações, sem caracterização de perfil e/ou etiquetamento social (*profiling*).

Repise-se, assim, que vislumbrar o *Big Data* como um direito coletivo autônomo advindo da normatividade de um bem jurídico transindividual, notadamente do meio ambiente digital, é fundamental para estabelecer textos legais capazes de nos oferecer a construção de normatividades principiológicas que vão além dos textos existentes, mas alcançam a realidade existencial por meio da interpretação integrativa.

Conclui-se que o *Big Data* na sua expressão transindividual está, portanto, também sujeito à transparência e à apurada identificação das fontes e dos controladores, a quem competem as decisões referentes ao tratamento de dados pessoais (art. 5º, inciso VI, LGPD) (BRASIL, 2018), ainda que visto como realidade autônoma e, provavelmente, alienado e conjugado como ente indissociável. Dá-se por certa a necessidade de se atribuir às respectivas responsabilidades quando do uso de seus conceitos, a fim de evitar abusos e assegurar um mínimo de equilíbrio do bem jurídico ambiental, ou seja, do meio ambiente digital.

5 Conclusão

Buscou-se apresentar as características do tratamento de dados realizados pelas empresas provedoras de aplicação, a exemplo das redes sociais, elencando o conceito e os aspectos técnicos de *Big Data*, como sendo o poder de tratamento e análise de dados semiestruturados ou não estruturados provenientes de diversas fontes, formatos, estruturas, origens e tipos, relevando-se como uma forma de coleta de dados que vai além de um mero banco de dados, pois armazena dados contextualizados, criando e moldando tendências em uma perene, complexa e dinâmica organização das relações sociais. Isto faz com que esse volume de dados se desvincule de sua origem e agregue característica de realidade autônoma aos usuários individuais, não só se propondo a elencar as preferências dos usuários, mas chegando ao esforço iniludível de moldar, modificar e até mesmo induzir opiniões, crenças e comportamentos àqueles expostos às suas práticas.

Visualizou-se o *Big Data* como um instrumento capaz de desequilibrar os pressupostos da economia de livre mercado, gerando falhas nas relações econômicas a ponto de mitigar a livre concorrência e, por consequência, violar direitos coletivos e transindividuais. Buscou-se explorar e emoldurar as possíveis falhas de mercado oriundas da livre disposição e utilização do *Big Data* pelas empresas provedoras de aplicação, o que permitiu chegar ao motivo das mitigações acentuadas ao bem-estar coletivo e aos interesses da sociedade com o uso desregulado e indiscriminado do *Big Data* por tais empresas.

Por fim, foram elencadas razões aliadas da necessidade de estruturação e proteção ampla do meio ambiente digital, com a necessidade de regulação da atividade econômica de tratamento de dados pelas empresas provedoras de aplicação, a exemplo das redes sociais, no intuito de diminuir ou eliminar essas falhas no desenvolvimento da economia, na tentativa de maximizar o bem-estar social coletivo e proteger os interesses metaindividuais, em regra indivisíveis e direcionados a destinatários indeterminados, enunciando a necessidade de se promover uma estrita regulação econômica das atividades de exploração do *Big Data* no mercado de consumo, notadamente com as diretrizes gerais e os princípios enunciativos da recentemente sancionada Lei Geral de Proteção dos Dados Pessoais.

Referências

AB2L. Premissas para a reflexão sobre a regulação da tecnologia, **Associação Brasileira de Lawtechs e Legaltechs**, 2018. Disponível em: <https://ab2l.org.br/premissas-para-a-reflexao-sobre-a-regulacao-da-tecnologia/>. Acesso em: 14 fev. 2019.

BALLART, Xavier; RAMIÓ, Carles Matas. **Ciencia de la administración pública**. Valência: Tirant Lo Blanch, 2002. (Colección Ciencia Política, v. 8).

BRASIL. Presidência da República. **Lei 13.709 de 14 de agosto de 2018**. Lei Geral de Proteção de Dados Pessoais. Brasília, DF: Presidência da República, 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2018/Lei/L13709.htm. Acesso em: 14 fev. 2019.

BRASIL. Presidência da República. **Lei 12.965 de 23 de abril de 2014**. Estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil. Brasília, DF: Casa Civil, 2014. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm. Acesso em: 14 fev. 2019.

BRASIL. [Constituição (1988)]. **Constituição da República Federativa do Brasil**. Brasília, DF: Casa Civil, 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 14 fev. 2019.

- BHARGAVA, Rahul *et al.* Beyond data literacy: reinventing community engagement and empowerment in the age of data; **Data-pop Alliance**, Nova Iorque, 29 set. 2015. White paper series. Disponível em: <https://datapopalliance.org/item/beyond-data-literacy-reinventing-community-engagement-and-empowerment-in-the-age-of-data/>. Acesso em: 04 mar. 2021.
- CAVEDON, Ricardo. **Teoria geral dos direitos coletivos**: releitura da racionalidade dos direitos fundamentais de terceira geração. Curitiba: Juruá, 2015.
- CAVEDON, Ricardo; FERREIRA, Helene Sivini; FREITAS, Cinthia Obladen de Almendra. O meio ambiente digital sob a ótica da teoria da sociedade de risco: os avanços da informática em debate. **Revista Direito Ambiental e Sociedade**, Caxias do Sul, v. 5, p. 194-223, 2015.
- FACEBOOK. **Política de dados**, 2016a. Disponível em: <https://www.facebook.com/privacy/explanation#>. Acesso em: 14 fev. 2019.
- FACEBOOK. **Política de Privacidade**, 2016b. Disponível em: https://www.facebook.com/legal/FB_Work_Privacy. Acesso em: 14 fev. 2019.
- FERRARIS, Valeria *et al.* Working paper: defining profiling. **United Nations Interregional Crime and Justice Research Institute (UNICRI)**, dez. 2013. Disponível em: www.unicri.it/news/files/Profiling_final_report_2014.pdf. Acesso em: 14 fev. 2019.
- FRANÇA, Tiago Cruz *et al.* Big Social Data: Princípios sobre coleta, tratamento e análise de dados sociais. *In*: SIMPÓSIO BRASILEIRO DE BANCO DE DADOS, 29, 2014, Curitiba. **Anais [...]**. Curitiba: UFPR, 2014. p. 8 - 45. Disponível em: <http://www.inf.ufpr.br/sbbd-sbsc2014/sbbd/proceedings/artigos/pdfs/127.pdf>. Acesso em: 14 fev. 2019.
- GANTZ, John; REINSEL, David. The digital universe in 2020: Big Data, nigger digital shadows, and biggest growth in the far east. **IDC**, New York, v. 3, Dec. 2012. Disponível em: <https://www.emc.com/leadership/digital-universe/2012iview/index.htm>. Acesso em: 14 fev. 2019.
- GRAU, Eros Roberto. **A ordem econômica na constituição de 1988**: (interpretação e crítica). São Paulo: Malheiros, 2008.
- GREENFIELD, Adam. **Everyware**: the dawning age of ubiquitous computing. San Francisco: New Riders, 2006.
- KITCHIN, Rob. **The data revolution**: big data, open data, data infrastructures & their consequences. Los Angeles: SAGE Publications, 2014.
- KUECHLER, William L. Business applications of unstructured text. **Communications of the ACM**, New York, v. 50, n. 10, 2007, p. 86-93. DOI: <https://doi.org/10.1145/1290958.1290967>. Acesso em: 14 fev. 2019.
- LAUDON, Kenneth C.; LAUDON, Jane Price. **Sistemas de informação**. Rio de Janeiro: Livros Técnicos e Científicos S.A., 1999.
- MANYIKA, James *et al.* **Big data**: the next frontier for innovation, competition, and productivity. [S.]: McKinsey Global Institute, 2011.
- MARX, Gary T. What's new about the "new surveillance"? Classifying for Change and Continuity. **Knowledge, Technology and Policy**, New York, v. 17, n.1, 2004. p. 18-37. Disponível em: <https://doi.org/10.1007/BF02687074>. Acesso em: 14 fev. 2019.
- MICHELL, William C.; SIMMONS, Randy T. **Para além da política**: mercados, bem-estar social e o fracasso da burocracia. Tradução de Jorge Ritter. Rio de Janeiro: Topbooks, 2003.
- MOUTINHO, Ana Maria. **Inteligência Ambiente**: contributo para a conceptualização de parede inteligente. 2010. Dissertação (Mestrado em Arte e Multimédia) - Universidade de Lisboa, Lisboa, 2010. Disponível em: http://repositorio.ul.pt/bitstream/10451/7277/2/ULFBA_tes%20392.pdf. Acesso em: 14 fev. 2019.

OPEN DATA HANDBOOK. **O que são dados abertos?**, 2017. Guide. Disponível em: http://opendatahandbook.org/guide/pt_BR/what-is-open-data/. Acesso em: 14 fev. 2019.

PARISER, Eli. **O Filtro Invisível**: o que a internet está escondendo de você. Tradução de Diego Alfaro. Rio de Janeiro: Zahar, 2012.

PEREIRA, Orlando P. Importância da Intervenção Pública na Economia de Mercado. **Estudios Económicos de Desarrollo Internacional - AEEADE**, [S.], v. 5, n. 1, p. 117-132, 2005. Disponível em: <http://www.usc.es/economet/reviews/eedi515.pdf>. Acesso em: 14 fev. 2019.

SCHWAB, Klaus. **A quarta revolução industrial**. Tradução de Daniel Moreira Miranda. São Paulo: Edipro, 2016.

SCHNEIDER, Christie. The biggest data challenges that you might not even know you have. **IBM**, 2016. Disponível em: <https://www.ibm.com/blogs/watson/2016/05/biggest-data-challenges-might-not-even-know/>. Acesso em: 14 fev. 2019.

SMITH, Adam. **A riqueza das nações**: (compêndio). Tradução de Bento da Silva. Lisboa: Editora Hunter Books, 2014.

UNIÃO EUROPEIA. **Regulation (EU) 2016/679**. General Data Protection Regulation – GDPR. 2016. Disponível em: <https://gdpr-info.eu/>. Acesso em: 14 fev. 20219.

VAZAMENTO em banco de dados expõe CPF de 220 milhões de pessoas. **Jornal Extra**. 17 de janeiro de 2021. Disponível em: <https://extra.globo.com/economia/vazamento-em-banco-de-dados-expoe-cpf-de-220-milhoes-de-pessoas-24845669.html>. Acesso em: 27 jan. 2021.

WEISER, Mark. The Computer for the 21st Century. **Scientific American**, EUA, p. 94-104, set., 1991. Disponível em: <https://www.lri.fr/~mbl/Stanford/CS477/papers/Weiser-SciAm.pdf> Acesso em: 14. fev. 2019.

WEISER, Mark. Some Computer Science Issues in Ubiquitous Computing. **Communications of the ACM**, New York, p. 75-84, jul. 1993. Disponível em: http://cseweb.ucsd.edu/~schulman/class/cse291_f18/docs/weiser_cacm.pdf. Acesso em: 14 fev. 2019.

Recebido em: 10/02/2020

Aprovado em: 26/01/2021