

# O MEIO AMBIENTE DIGITAL E O DIREITO À PRIVACIDADE DIANTE DO *BIG DATA*

Émilien Vilas Boas Reis<sup>1</sup>

Escola Superior Dom Helder Câmara (ESDHC) |

Bruno Torquato de Oliveira Naves<sup>2</sup>

Escola Superior Dom Helder Câmara (ESDHC) |

## RESUMO

Convive-se, hoje, com um grande volume de dados – *Big Data* –, o que representa uma revolução nos moldes de comercialização, propaganda, concorrência e, até mesmo, na própria delimitação do espaço privado. A privacidade no meio ambiente digital deve ser considerada sob a possibilidade de consequências muito mais gravosas, no caso de violação. Assim, este artigo propõe-se a analisar a regulação jurídica do meio ambiente digital, sobretudo diante do direito à privacidade. Para tanto, realizou-se pesquisa bibliográfica (teórico-qualitativa), com conclusões argumentativas que permitiram perceber o direito à privacidade com maior fluidez diante dos dados pessoais, aproximando-o dos princípios jurídicos. Analisou-se a regulamentação realizada pela Lei Geral de Proteção de Dados Pessoais, em especial a ambivalência dos dados pessoais sensíveis.

**Palavras-chave:** *Big Data*; dados sensíveis; Lei Geral de Proteção de Dados Pessoais; meio ambiente digital; privacidade.

1 Pós-doutor em filosofia pela Universidade do Porto (UP). Doutor em filosofia pela Pontifícia Universidade Católica do Rio Grande do Sul (PUC-RS). Mestre em Filosofia pela PPUC-RS. Graduado em Filosofia pela Universidade Federal de Minas Gerais (UFMG). Professor adjunto da ESDHC em nível de graduação e pós-graduação (mestrado e doutorado). ORCID: <http://orcid.org/0000-0003-0729-522X> / e-mail: [mboasr@yahoo.com.br](mailto:mboasr@yahoo.com.br)

2 Doutor e Mestre em Direito pela Pontifícia Universidade Católica de Minas Gerais (PUC-MINAS). Professor nos cursos de mestrado e doutorado em Direito Ambiental e Desenvolvimento Sustentável da ESDHC. Coordenador do curso de especialização em Direito Urbanístico e Ambiental da PUC-MINAS Virtual. Professor na PUC-MINAS. Pesquisador do Centro de Estudos em Bioética (CEBID). ORCID: <https://orcid.org/0000-0001-8329-9420> / e-mail: [brunotorquato@hotmail.com](mailto:brunotorquato@hotmail.com)

*THE DIGITAL ENVIRONMENT AND THE RIGHT TO PRIVACY IN RELATION TO BIG DATA*

*ABSTRACT*

*Today, we live with a large volume of data – Big Data –, which represents a revolution in the forms of commercialization, advertising, competition and even in the very delimitation of private space. Privacy in digital environment should be considered under the possibility of much more serious consequences in case of violation. Thus, this paper suggests analyzing the legal regulation of the digital environment, especially in relation to the right to privacy. Therefore, a bibliographic research (theoretical-qualitative) was carried out, with argumentative conclusions that allowed to perceive the right to privacy with greater fluidity considering personal data, bringing it closer to legal principles. The regulation made by the General Law on the Protection of Personal Data was analyzed, particularly the ambivalence of sensitive personal data.*

**Keywords:** *Big data; digital environment; General Law on Protection of Personal Data; privacy; sensitive data.*

## INTRODUÇÃO

Para além de qualquer concepção platônica, vive-se, atualmente, em dois mundos: o mundo concreto e o mundo digital. Nos dois mundos, o indivíduo nasce, se constrói e estabelece relacionamentos. No mundo digital, no entanto, não se morre.

Os dados digitais têm persistência histórica. Ao que parece, enquanto existir humanidade, eles permanecerão. Fala-se até de testamento de bens digitais para tratar da destinação jurídica de manifestações voluntárias em mídias sociais, que acumulam dados pessoais e têm valor intangível, como contas de *Facebook*, *Instagram* e *Twitter*.

O grande volume de informações disponíveis digitalmente é o que se denomina de *Big Data*. Um conjunto de dados de todas as categorias e formatos compõe um universo de múltiplos interesses. Mas como será que o Direito está pronto para mediar todos esses interesses, por vezes contrapostos? A regulação jurídica do meio ambiente digital é hoje uma preocupação de diversos países.

Há um movimento mundial em torno do tráfego de dados, visando determinar formas aceitáveis de tratamento, controle e compartilhamento dessas informações. Em 2016, a União Europeia aprovou o Regulamento n. 679, mais conhecido como Regulamento Geral sobre a Proteção de Dados, que objetivou a unificação da política de dados entre os países. O Brasil, mais recentemente, ingressou nesse movimento com a Lei Geral de Proteção de Dados Pessoais.

Certo é que a tecnologia de dados está desafiando normas jurídicas e visões pré-estabelecidas, o que reflete na transformação da própria concepção da vida privada do ser humano.

Este artigo pretende analisar como fica a privacidade nesse quadro em que se avolumam informações, mas cuja circulação nem sempre se dá de maneira intencional. Trata-se de pesquisa teórico-qualitativa, empreendida por meio de fontes bibliográficas que contribuíssem para questões éticas e jurídicas trazidas pela relação entre *Big Data* e inteligência artificial (IA), a fim de investigar as novas conformações do direito à privacidade.

### 1 CONSIDERAÇÕES ACERCA DO *BIG DATA*

A humanidade está vivendo uma grande revolução no que se refere às informações e seus usos. Em nenhuma outra época se produziu tantos dados, que estão se acumulando na internet.

A cada segundo, o Google processa mais de 40.000 pesquisas; totalizando 3,5 bilhões por dia. A cada minuto, usuários do Snapchat compartilham 527.760 fotos e usuários do YouTube assistem a mais de 4,1 milhões de vídeos. Há ainda os sistemas antigos, como emails, que continuam a experimentar um crescimento significativo. A cada minuto, 156 milhões de mensagens são enviadas.

No entanto, outro aspecto precisa ser considerado: empresas e máquinas também geram enormes quantidades de dados. De acordo com uma pesquisa do Statista, o número de sensores alcançará 12,86 bilhões até 2020.

À luz de tudo isso, parece ser uma boa aposta que os volumes de dados continuarão a aumentar rapidamente. Em um relatório da International Data Corporation (IDC) chamado “Data Age 2025” (“Data 2025”), espera-se que a quantidade de dados criados até 2025 alcance surpreendentes 163 zettabytes. Isso é aproximadamente dez vezes a quantidade de 2017 (TAULLI, 2020).

Após a invenção do computador, das linguagens computacionais, da internet e das redes sociais, o mundo se depara com um novo termo para ilustrar a capacidade de lidar com esse acúmulo de dados sem precedentes. A expressão é *Big Data*. A ideia por trás do termo é de que as infindáveis informações disponíveis e acumuladas, e hoje presentes no universo *on-line*, serão úteis para várias áreas.

Things really are speeding up. The amount of stored information grows four times faster than the world economy, while the processing power of computers grows nine times faster. Little wonder that people complain of information overload. Everyone is whiplashed by the changes (MAYER-SCHONBERGER; CUKIER, 2017).

Um dos exemplos mais intrigantes da capacidade do *Big Data* foi relatado no texto *Detecting influenza epidemics using search engine query data*, em que pesquisadores do Google (GINSBERG et al., 2009), a partir de consultas de pesquisa do próprio Google, buscaram rastrear o influenza (H1N1) nos Estados Unidos.

Os sistemas de vigilância de doenças, como o Centro de Controle e Prevenção de Doenças dos Estados Unidos (*Centers for Disease Control and Prevention – CDC*) e o Sistema Europeu de Vigilância da Gripe (*European Influenza Surveillance Scheme – EISS*), fazem relatórios de controle que não saem com resultados imediatos. Na tentativa de acelerar esse processo, os autores usaram como método de pesquisa a análise dos dados de consulta do Google relacionados à influenza durante cinco anos. Os dados coletados compreenderam os anos de 2003 e 2008 e levaram em conta os 50 milhões de termos mais consultados nos Estados Unidos, durante o período de maior avanço da doença.

Ginsberg et al. (2009) pegaram os termos mais procurados no Google

e compararam com os dados do CDC sobre a propagação de gripe em cada uma das nove regiões divididas pelo CDC e com as consultas ambulatoriais. Os autores estavam tentando criar um modelo de previsão a partir da relação das pesquisas no Google com os demais dados disponíveis. Ao longo da temporada de gripe entre 2007 e 2008, usando seu modelo de previsão, os pesquisadores conseguiram estimar a propagação da gripe de uma a duas semanas antes do relatório oficial do CDC. A constatação fora de que as consultas de pesquisa no Google podiam “antecipar” e “prever” a propagação da gripe antes dos órgãos oficiais e antes que houvesse uma epidemia. Em suma:

Harnessing the collective intelligence of millions of users, Google web search logs can provide one of the most timely, broad-reaching influenza monitoring systems available today. Whereas traditional systems require 1–2 weeks to gather and process surveillance data, our estimates are current each day (GINSBERG et al., 2009, p. 1014).

O número de dados (informações) tem crescido exponencialmente a cada ano. O *Big Data* utiliza-os para fazer previsões. Pode-se compreender os dados em: (a) *bit* (forma curta de *Binary digit* – dígito binário), que é expresso em 0 ou 1 e é usado para medir a quantidade de dados na rede; (b) *byte* (que significa o armazenamento – p.ex. um arquivo tem *X Megabytes*) (TAULLI, 2020). Por outro lado, os dados têm diferentes fontes de origens:

- Web/rede social (Facebook, Twitter, Instagram, YouTube)
- Dados biométricos (rastreadores de atividades físicas, testes genéticos)
- Sistemas de ponto de venda (de lojas físicas e sites de comércio eletrônico)
- Internet das coisas ou IoT (Etiquetas de identificação e dispositivos inteligentes)
- Sistemas de nuvem (aplicativos de negócios como Salesforce.com)
- Bancos de dados corporativos e planilhas (TAULLI, 2020).

Ainda é possível identificar os dados a partir de sua organização: (a) dados estruturados, que são armazenados em bancos de dados ou planilhas e têm pequeno volume (p. ex.: informações financeiras, endereços, informações sobre produtos, números de telefone.); (b) dados não estruturados, que não possuem formatação predefinida (p. ex.: imagens, vídeos, arquivos de áudio e de texto, informações de redes sociais etc.); (c) dados semiestruturados, que contam dados estruturados e não estruturados (p. ex.: XML – *Extensible Markup Language*, JSON – *JavaScript Object Notation* –, que são interfaces de programação de aplicativo); e (d) dados temporais, que podem ser estruturados, não estruturados e semiestruturados e são informações interativas (p. ex.: a junção de informações sobre um cliente, que

vai a um *site* ou usa um aplicativo) (TAULLI, 2020).

Em 2001, Doug Laney, da empresa de consultoria Gartner, escreveu o texto *3D Data Management: Controlling Data Volume, Velocity, and Variety*, em que sugeriu que os dados deveriam ser analisados a partir do volume, da velocidade e de sua variedade. Sobre o volume, Laney já alertava sobre o aumento considerado de dados, o que significava que a velocidade também deveria aumentar, a fim de lidar com a grande quantidade de dados. Por fim, a variedade dos dados que, devido aos diferentes tipos de fontes de informação, também deveria ser levada em consideração (LANEY, 2001).

Apesar de ter sido um pioneiro, analistas sobre os dados afirmavam que a proposta de Laney era insuficiente. Por isso, propuseram outros elementos que deveriam ser levados em consideração no que diz respeito aos dados.

A seguir, estão alguns dos mais comuns:

- Veracidade: refere-se a dados considerados precisos. Neste capítulo, vamos discutir algumas das técnicas para avaliar a veracidade.
- Valor: mostra a utilidade dos dados. Muitas vezes, refere-se a ter uma fonte confiável.
- Variabilidade: significa que os dados geralmente mudam ao longo do tempo. É o caso, por exemplo, do conteúdo de mídia social que pode se transformar com base no sentimento geral com relação a novos acontecimentos e notícias de última hora.
- Visualização: trata do uso de recursos visuais – como gráficos – para entender melhor os dados (TAULLI, 2020).

*Big Data* talvez tenha três grandes características: (a) possibilita a análise de grandes quantidades de dados; (b) tem a pretensão de organizar o mundo real; e (c) lida com correlação e não com causalidades (MAYER-SCHONBERGER; CUKIER, 2017).

O exemplo trazido sobre pesquisadores do Google e as gripes sazonais pode ilustrar isso. Foram analisados mais de 50 milhões de palavras pesquisadas (incidência, localidade, hora etc.) e, a partir disso, os pesquisadores buscaram organizar o caos de informações, sem se preocuparem porque a gripe ocorria. A análise de dados depende da capacidade de coletar, armazenar e processar os dados. E uma vez que os dados podem ser recolhidos de maneira passiva, os vieses podem ir diminuindo.

A grande quantidade de dados analisados é o que torna o *Big Data* diferenciado do que se teve com manipulação de informações até então. O avanço tecnológico é importante, assim como melhores algoritmos, mas é a imensa massa de informação que vem tornando o *Big Data* mais

previsível a cada dia. Por exemplo, computadores passaram a vencer os seres humanos no jogo de xadrez, que tem as mesmas regras há séculos, porque o sistema se alimentava de mais dados com o passar dos anos (MAYER-SCHONBERGER; CUKIER, 2017).

O caos da realidade é abrangido pelo *Big Data*, na medida em que apenas 5% dos dados digitais são estruturados. A grande quantidade de dados faz que os 95% não estruturados também sejam apreendidos:

According to some estimates only 5 percent of all digital data is “structured” – that is, in a form that fits neatly into a traditional database. Without accepting messiness, the remaining 95 percent of unstructured data, such as web pages and videos, remain dark. By allowing for imprecision, we open a window into an untapped universe of insights (MAYER-SCHONBERGER; CUKIER, 2017).

Os resultados alcançados pelo *Big Data* não estão relacionados com o motivo e a causalidade. Os dados são usados em busca de correlações e probabilidades. Caminha-se para a possibilidade de prever eventos que ainda não ocorreram. O próprio método científico que começa com hipóteses para serem comprovadas corre risco de ser questionado.

By letting us identify a really good proxy for a phenomenon, correlations help us capture the present and predict the future: if A often takes place together with B, we need to watch out for B to predict that A will happen. Using B as a proxy helps us capture what is probably taking place with A, even if we can’t measure or observe A directly. Importantly, it also helps us predict what may happen to A in the future. Of course, correlations cannot foretell the future, they can only predict it with a certain likelihood. But that ability is extremely valuable (MAYER-SCHONBERGER; CUKIER, 2017).

A abrangência do *Big Data* também passa pela capacidade atual de quantificar os fenômenos e pelo fato de eles estarem inseridos no mundo *on-line*. Mais do que nunca, a frase de Galileu, de que Deus escreve o mundo em caracteres matemáticos, se mostra acertada. Livros, geolocalizadores, relações sociais, tudo está sendo transformado em dados. Pesquisadores da Cornell University analisaram em uma escala global o comportamento dos indivíduos em diferentes países e culturas a partir dos dados coletados da mídia social Twitter. Os autores tiveram como dados 509 milhões de mensagens de 2,4 milhões de pessoas em todo o mundo:

Using Twitter.com’s data access protocol, we collected up to 400 public messages from each user in the sample, excluding users with fewer than 25 messages. The resulting corpus contained about 2.4 million individuals from across the globe and 509 million messages authored between February 2008 and January 2010 (GOLDER; MAC, 2011, p. 1879-1880).

Os pesquisadores identificaram algumas características comuns entre os pesquisados: os indivíduos acordam com bom humor, mas tal sensação vai mudando com o passar do dia, o que se relaciona com o ritmo circadiano e com o sono. Também constataram que as pessoas são mais felizes nos finais de semana, e que, por acordarem mais tarde nesses dias, a maior afetividade positiva ocorre duas horas mais tarde do que durante a semana (GOLDER; MAC, 2011).

Com o *Big Data*, o mundo passa a ser visto como uma série de informações que devem ser trabalhadas como dados, inclusive pelas ciências humanas.

Outro ponto importante a se destacar é que os dados não são fornecidos apenas voluntariamente. Muitos dados são captados sem a anuência de seus titulares, utilizando-se de meios desconhecidos por estes:

Data can frequently be collected passively, without much effort or even awareness on the part of those being recorded. And because the cost of storage has fallen so much, it is easier to justify keeping data than discarding it. [...] Websites log every click users make – sometimes even where the mouse-cursor moves – for analyzing and optimizing the content the sites present to visitors (MAYER-SCHONBERGER; CUKIER, 2017).

*Sites* hoje têm a capacidade de registrarem os cliques dos usuários, assim como guardarem informações apenas por meio do movimento que se faz no *mouse*. No mundo *on-line* qualquer ação do usuário tem a capacidade de registrar dados, e, mesmo que a princípio aparentem ser inúteis, no infindável armazenamento de dados e por meio de relações com outros dados, passam a ser úteis. Os teóricos chamam os rastros deixados na internet de “exaustão de dados”:

A term of art has emerged to describe the digital trail that people leave in their wake: “data exhaust”. It refers to data that is shed as a byproduct of people’s actions and movements in the world. For the Internet, it describes users’ online interactions: where they click, how long they look at a page, where the mouse-cursor hovers, what they type, and more (MAYER-SCHONBERGER; CUKIER, 2017).

Mas existem dados que são coletados a partir do mundo *off-line*, como, por exemplo, quando varejistas, por meio de câmeras, rastreiam seus clientes em lojas físicas, captando até para onde os clientes estão olhando. De posse dessas informações, os lojistas podem manipular os dados a favor de melhores campanhas de *marketing* (MAYER-SCHONBERGER; CUKIER, 2017).

Dados que antes eram colhidos por um motivo, ao serem armazenados,

podem ser reutilizados para outros objetivos futuros. Grandes quantidades de dados não são descartadas simplesmente. Uma das perguntas que se pode fazer com o *Big Data* é: na medida em que os dados são precisos sobre as coisas, como ficam a privacidade e a liberdade humanas?

Grandes empresas privadas e os governos estão de posse desses grandes dados:

The Internet has made tracking easier, cheaper, and more useful. And clandestine three-letter government agencies are not the only ones spying on us. Amazon monitors our shopping preferences and Google our browsing habits, while Twitter knows what's on our minds. Facebook seems to catch all that information too, along with our social relationships. Mobile operators know not only whom we talk to, but who is nearby (MAYER-SCHONBERGER; CUKIER, 2017).

O acúmulo de informações pessoais leva ao questionamento de como, com qual interesse e por quem esses dados serão utilizados. São questões que o *Big Data* coloca e que talvez não possam ser mais evitadas, dado que há uma tendência mundial em se aumentar o gasto com dados:

A quantidade de dinheiro investida em dados é enorme. De acordo com a IDC (International Data Corporation –Corporação Internacional de Dados), prevê-se que os gastos com soluções de Big Data e analytics passarão de US\$ 166 bilhões em 2018 para US\$ 260 bilhões até 2022. Isso representa uma taxa de crescimento anual de 11,9%. Os maiores investidores incluem bancos, fabricantes discretos e de processos, empresas de serviços profissionais e Governo Federal. Eles respondem por quase metade do investimento total (TAULLI, 2020).

Os dados, por sua vez, dependem de uma outra área fundamental para sua utilização: inteligência artificial.

## 2 CONSIDERAÇÕES ACERCA DA INTELIGÊNCIA ARTIFICIAL (IA)

Apesar de estar no imaginário popular a automatização completa de robôs, que tomarão o mundo para si e que destruirão ou farão de escrava a humanidade, a inteligência artificial ainda não condiz com tal descrição. Entretanto, é outra parte importante para se compreender o uso de dados.

A IA tem uma longa história e seu precursor é Alan Turing (1912-1954)<sup>3</sup>. O impulso da IA atual, por sua vez, dependeu dos seguintes fatores: 1) crescimento explosivo de *datasets* (conjunto de dados): elemento trabalhado no tópico anterior; 2) infraestrutura:

<sup>3</sup> Para uma breve retomada do percurso histórico da IA até os dias atuais, cf. TAULLI, 2020.

que foi impulsionado pelo Google nos últimos 15 anos; 3) Unidades de processamento gráficos (GPUs – *Graphics Processing Units*), tecnologia de chips criados pela empresa de tecnologia NVIDIA (TAULLI, 2020).

Entretanto, há outros elementos a serem considerados ao se falar em IA. Um primeiro termo é o *machine learning*. Em 1959, Arthur L. Samuel publicou o artigo *Some Studies in Machine Learning Using the Game of Checkers*. Ele, que trabalhava na IBM, colocava em artigo o que já fizera de maneira prática como programador de um jogo de damas para computadores: a noção de como uma máquina, a partir de algumas regras e de estatística avançada, teria a capacidade de “aprender” algo, fazendo que se aperfeiçoasse. Em suas palavras:

Enough work has been done to verify the fact that a computer can be programmed so that it will learn to play a better game of checkers than can be played by the person who wrote the program. Further-more, it can learn to do this in a remarkably short period of time (8 or 10 hours of machine-playing time) when given only the rules of the game, a sense of direction, and a redundant and incomplete list of parameters which are thought to have something to do with the game, but whose correct signs and relative weights are unknown and unspecified. The principles of machine learning verified by these experiments are, of course, applicable to many other situations (SAMUEL, 1983, p. 211).

A linguagem que o *machine learning* utiliza é o algoritmo, que são procedimentos para se executar algo. Existem inúmeros, mas que podem ser abarcados em 4 categorias: (1) aprendizagem supervisionada: lida com dados rotulados, por exemplo, uma foto identificada; (2) aprendizagem não supervisionada: lida com dados não rotulados, assim, buscam-se padrões nos dados, por exemplo, grupos de clientes; (3) aprendizagem por reforço: processo que ocorre por tentativa e erro, tornando o aprendizado melhor com o passar do tempo e tendo mostrado-se fundamental para a melhora da IA; (4) aprendizagem semissupervisionada: leva em consideração a linguagem supervisionada e a linguagem não supervisionada. Nesse caso, a linguagem não supervisionada é transformada em linguagem supervisionada, por meio de *deep learning* (TAULLI, 2020).

A propósito, *deep learning* é uma área do *machine learning*. É um sistema que processa inúmeros dados com o objetivo de encontrar determinados padrões. *Deep learning* tem sido tomado como o elemento fundamental para os avanços da IA. Uma das fontes de inspiração para a *deep learning* é a comparação de que o cérebro é uma máquina (computador) e, portanto, pensar o sistema computacional à luz das funções cerebrais faria todo o sentido (TAULLI, 2020).

A estrutura básica para um modelo de *deep learning* é a denominada Rede Neural Artificial (*Artificial Neural Network* – ANN):

[...] é uma função que inclui unidades (que também podem ser chamadas de neurônios, perceptrons ou nós). Cada unidade terá um valor e um peso, que indicam sua importância relativa, e irá para a camada oculta. A camada oculta usa uma função, cujo resultado se torna a saída (TAULLI, 2020).<sup>4</sup>

Uma das técnicas utilizadas para o aperfeiçoamento da *deep learning* é a retroprogramação, que consiste em melhorar a rede neural a partir dos erros encontrados, inserindo novos valores na rede (TAULLI, 2020).

Existem diferentes tipos de redes neurais, destacando-se: (1) redes neurais conectadas – é a rede mais básica, onde há conexões entre os neurônios; (2) redes neurais recorrentes (*recurrent neural network* – RNN) – é uma rede que, além de processar a entrada atual, processa também as entradas anteriores, por exemplo, buscador automático de palavras; (3) redes neurais convolucionais (*convolutional neural network* – CNN) – é uma rede que leva em consideração diferentes convoluções (variações) de análise de dados, agregando e combinando os diferentes dados, por exemplo, o reconhecimento de faces; (4) redes adversárias generativas (GANs – *generative adversarial network*) – é uma rede que cria novos tipos de saída (p. ex. vídeos e áudios) a partir de dados introduzidos (TAULLI, 2020).

Outro item importante quando se fala em IA é a Automação Robótica de Processos (RPA – *Robotic Process Automation*), que não está relacionada a robôs físicos, mas a robôs softwares. Os *bots* (diminutivo de robô) reproduzem um trabalho repetitivo (TAULLI, 2020).

Assim, apresentados às noções de *Big Data* e IA, o próximo item se debruçará sobre os desafios éticos implicados.

### 3 DESAFIOS ÉTICOS E JURÍDICOS EM RELAÇÃO AO *BIG DATA*, ALGORITMOS E INTELIGÊNCIA ARTIFICIAL

O mundo contemporâneo traz problemas práticos que forcem diferentes reflexões em relação às novas descobertas. A junção de *Big Data*, Inteligência Artificial e algoritmos trazem questões difíceis de serem resolvidas. Propositamente, os primeiros capítulos evitaram enfatizar tais problemas, mas que agora se tornam oportunos.

O uso dos dados, muitas vezes sem a convivência dos envolvidos, a dependência mediante os algoritmos, a exclusão do trabalho humano ao

<sup>4</sup> Camada oculta é o nível de análise do *deep learning*.

longo desses novos processos tecnológicos, o fim de determinadas carreiras, o questionamento sobre a existência de uma vontade livre, a dificuldade do direito em lidar com os novos desafios e a diminuição da responsabilidade são algumas das questões envolvidas.

Um ponto interessante que os tempos atuais trazem é que a reflexão ética não deve ocorrer no que diz respeito aos aparelhos em si (*hardware*), mas ao uso dos dados, aos algoritmos programados e à inteligência artificial, portanto, no nível de *softwares*:

It is not the hardware that causes ethical problems, it is what the hardware does with the software and the data that represents the source of our new difficulties. LoAD brings into focus the different moral dimensions of data. In doing so, it highlights the fact that, before concerning information, ethical problems such as privacy, anonymity, transparency, trust and responsibility concern data collection, curation, analysis and use, and hence they are better understood at that level (FLORIDI; TADDEO, 2016, p. 3).<sup>5</sup>

Luciano Florido, um dos proeminentes estudiosos sobre ética dos dados, afirma que tal ramo da ética estuda problemas morais relacionados aos dados, que incluem geração, registro, processamento, divulgação etc., mas também estudos sobre algoritmos, inteligência artificial e condutas sobre tais técnicas (FLORIDI; TADDEO, 2016).

A ética dos dados se debruça sobre a coleta e a análise de dados, o que inclui o recurso do *Big Data* em áreas tão distintas como as biomédicas, as ciências sociais, a publicidade e outras. Aspectos como a confiança e a transparência sobre os dados também são aspectos fundamentais:

In this context, key issues concern possible re-identification of individuals through data-mining, –linking, –merging and re-using of large datasets, as well as risks for so-called ‘group privacy’, when the identification of types of individuals, independently of the de-identification of each of them, may lead to serious ethical problems, from group discrimination (e.g. ageism, ethnicism, sexism) to group-targeted forms of violence (FLORIDI; TADDEO, 2016, p. 3)

A ética dos algoritmos é uma subárea que se torna fundamental devido ao aumento da capacidade dos algoritmos e por sua autonomia crescente, por meio da inteligência artificial e sua capacidade de aprendizado. Também inclui análise de situações indesejadas, como discriminação. “In this case, some crucial challenges include moral responsibility and accountability of both designers and data scientists with respect to unforeseen and undesired consequences as well as missed opportunities” (FLORIDI; TADDEO, 2016, p. 3).

<sup>5</sup> LoAD (*level of abstraction data-centric*)

A subárea ética prática (deontologia profissional) lida com a responsabilidade de indivíduos e empresas que se debruçam sobre os dados: “[...] with the goal to define an ethical framework to shape professional codes about responsible innovation, development and usage, which may ensure ethical practices fostering both the progress of data science and the protection of the rights of individuals and groups” (FLORIDI; TADDEO, 2016, p. 3).

Uma ética dos dados engloba a ética dos dados propriamente dita, mas também a ética dos algoritmos e a ética prática, o que a torna um estudo transdisciplinar.

#### **4 MEIO AMBIENTE DIGITAL E A FUNÇÃO PREDITIVA DOS ALGORITMOS**

No Direito Ambiental, entre as didáticas classificações de meio ambiente, pode-se hoje abrir uma nova categoria, atinente ao meio ambiente digital, onde novos tipos de relacionamentos interpessoais se travam, por vezes com pressupostos diferentes, a partir de dados voluntariamente fornecidos ou compulsória ou ilegalmente coletados.

Celso Antônio Pacheco Fiorillo coloca o meio ambiente digital como uma divisão do meio ambiental cultural:

O meio ambiente cultural, por via de consequência, manifesta-se no século XXI em nosso país exatamente em face de uma cultura que passa por diversos veículos reveladores de novo processo civilizatório adaptado necessariamente à sociedade da informação, a saber, de nova forma de viver relacionada a uma cultura de convergência em que as emissoras de rádio, televisão, o cinema, os *videogames*, a internet, as comunicações por meio de ligações de telefones fixos e celulares etc. moldam uma nova vida reveladora de nova faceta do meio ambiente cultural, a saber: o meio ambiente digital (FIORILLO, 2012, p. 81).

A Internet integra, sem dúvida, uma parte significativa do meio ambiente digital, mas há também dados *off-line* que são importantes componentes desse espaço, embora a tendência seja a comunicação de todos os dados por redes internas ou pela rede mundial de computadores.

Em razão da alta quantidade de informação no meio ambiente digital, o contato humano com esse ambiente se faz por meio de algoritmos, que são os instrumentos capazes de “minerar” essas informações, ou mesmo estabelecer correspondências entre elas, e devolver resultados inteligíveis a partir do critério utilizado de busca. Suas funções serão, basicamente,

descritiva ou preditiva, extraindo padrões representativos de informações.

A função algorítmica descritiva realiza a compilação e síntese dos dados, já a função preditiva almeja antever comportamentos por meio da análise de dados.

Exemplo interessante de função preditiva, a partir de dados pessoais, é narrado por Charles Duhigg (2012), em “O Poder do Hábito”.

Ao receber em casa cupons e folhetos de produtos para gestante, endereçados à sua filha adolescente, o pai irou-se e foi a uma Target, em Minnesota, tirar satisfação com o gerente. Como a menina ainda estava no ensino médio, seu questionamento era muito razoável, pois parecia que a Target queria influenciar sua filha a ter um bebê.

O gerente, constrangido, teve de pedir desculpas.

A Target, uma grande rede estadunidense, vende desde produtos de supermercado até artigos de loja de departamento. Ela queria descobrir os hábitos de cada comprador individual, a fim de direcionar propagandas personalizadas. Por isso, em 2002, contratou o estatístico Andrew Pole, para analisar as preferências de compras de cada consumidor.

Se você usa seu cartão de crédito da Target para comprar uma caixa de picolés uma vez por semana, geralmente por volta das seis e meia da noite, e sacos de lixo gigantes nos meses de julho e outubro, os estatísticos da Target e programas de computador determinam que você tem crianças em casa, tende a parar para comprar comida no caminho de volta do trabalho e tem um gramado que precisa ser aparado no verão, além de árvores que soltam folhas no outono. Vão examinar seus outros padrões de compra e notar que às vezes você compra cereal matinal, mas nunca leite – o que significa que deve estar comprando leite em algum outro lugar. Por isso a Target vai lhe enviar cupons de desconto no leite desnatado e também no chocolate granulado, no material escolar, em móveis para jardim, rastelos e – já que é provável que você queira relaxar após um longo dia de trabalho – cerveja. A empresa vai adivinhar o que você costuma comprar, e então tentar convencê-lo a comprar isso na Target. Ela tem a capacidade de personalizar os anúncios e cupons que envia para cada cliente, embora você talvez nunca perceba que recebeu pelo correio um folheto diferente do de seus vizinhos (DUHIGG, 2012, p. 231).

Um dos grandes nichos de mercado é o destinado a novos pais. Uma pesquisa da Target, em 2010, estimou que os pais gastam uma média de 6.800 dólares em produtos de bebê, antes de seu primeiro aniversário.

A partir da estatística de hábitos de compra, fornecidos por futuras mães – como seu nome, o nome do cônjuge e a data prevista para o nascimento – Pole detectou padrões de compras relativos a cada fase da gestação e desenvolveu um sistema de predição. Com uma combinação de

produtos adquiridos, Pole podia prever se uma mulher estava grávida e em que fase da gestação estava e assim direcionar propagandas e cupons.

Jenny Ward, uma mulher de 23 anos, de Atlanta, que comprou loção de manteiga de cacau, uma bolsa grande o bastante para servir como porta-fraldas, zinco, magnésio e um tapete azul? Há uma chance de 87% de que ela esteja grávida, e que seu parto esteja previsto para o final de agosto. Liz Alter, do Brooklyn, 35 anos, que comprou cinco pacotes de toalhas de mão, um frasco de sabão em pó para “pele sensível”, jeans folgados, vitaminas contendo DHA e um monte de hidratantes? Ela tem uma chance de 96% de estar grávida e provavelmente vai dar à luz no começo de maio. Caitlin Pike, uma mulher de 39 anos, de São Francisco, que comprou um carrinho de bebê de 250 dólares, mas nada além disso? Provavelmente está comprando para o chá de bebê de uma amiga. Além disso, seus dados demográficos mostram que ela se divorciou há dois anos (DUHIGG, 2012, p. 236).

E foi assim que a Target previu que aquela adolescente estava grávida. O gerente da loja de Minnesota, após alguns dias, ligou para o pai daquela adolescente para se desculpar novamente. E o pai, ainda mais constrangido, contou que tivera uma conversa com sua filha e descobrira que ela estava mesmo grávida.

Esse caso levanta a questão da privacidade de informações e das modificações pelas quais o mercado tem passado nos últimos anos, com o oferecimento de produtos “sob medida” para cada usuário.

## **5 DIREITO À PRIVACIDADE E À INTIMIDADE NO AMBIENTE DIGITAL**

Na Modernidade, a razão trouxe consigo a necessidade de conceituações precisas e classificações, mas também ressaltou que a única saída para o Direito era reconhecer a individualidade e o indivíduo como seu agente construtor.

Considerada como “aquilo que nos permite definir o que é e o que não é importante para nós” (TAYLOR, 1997, p. 47), a identidade permite que as potencialidades do indivíduo se concretizem segundo seus próprios interesses e convicções.

A pluralidade do homem e o projeto inacabado de construção de sua personalidade dependem das noções de intimidade e de privacidade, o que refletirá na autonomia como elemento determinante da dignidade do ser humano.

O desenvolvimento tecnológico é produto dessa autonomia humana

na técnica, mas, com ela, o próprio criador pode tornar-se refém de sua obra, em razão dos riscos a que se expõe.

Assim, o poder de autodeterminação do indivíduo deve também se voltar para o controle dos dados armazenados e a tomada de decisão após sua coleta.

Em se tratando de controle, tratamento e compartilhamento de dados, há de se observar o direito subjetivo à informação prévia e completa do procedimento a ser realizado. Essas informações compõem a própria autonomia, já que seu exercício requer, além do discernimento, consciência da situação envolvida, com as vantagens e os riscos envolvidos, especialmente quando se tratar de dados sensíveis.

Há dados, como os referentes à saúde, que após a coleta, deve-se questionar se até seu titular tem interesse em conhecê-los ou não.

Aliás, os dados de saúde somente poderão ser coletados, utilizados e conservados para os fins de diagnóstico e assistência sanitária; investigação médica, farmacológica e outras formas de investigação científica, como estudos de caráter antropológico; medicina forense e procedimentos civis ou penais ou outras atuações legais.

A questão da intimidade e da privacidade na revelação dos dados é de extrema relevância para a análise. Sabe-se que, sendo os dados armazenados, estes podem revelar preferências, hábitos, pensamentos, manifestações da imagem, do corpo e da saúde humanas.

Os direitos à intimidade e à privacidade são garantidos na Constituição Federal de 1988 por meio do disposto no inciso X<sup>6</sup> do artigo 5<sup>o</sup>. No Código Civil, a vida privada, em seu duplo aspecto, de intimidade e privacidade, está tutelada pelo art. 21, que assim dispõe: “A vida privada da pessoa natural é inviolável, e o juiz, a requerimento do interessado, adotará as providências necessárias para impedir ou fazer cessar ato contrário a esta norma”.

A intimidade é a esfera de projeção do indivíduo em sua relação interior. Já o direito à privacidade é um círculo pessoal maior, pois envolve as relações interpessoais do indivíduo. Enquanto a intimidade situa-se no compartimento mais restrito da pessoa, com situações que não se deseja compartilhar, a privacidade retrata a vida pública, familiar ou social do indivíduo, abrangendo o direito ao controle de coleta e utilização de dados pessoais.

6 “X – 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”.

O exercício do direito da personalidade à vida privada se faz de duas formas principais: a) um exercício positivo, consistente na livre condução da própria vida, a partir do compartilhamento de ideias e pensamentos, de exposição pública, inclusive com a disponibilidade, por meio de negócios jurídicos, de imagens, objetos e manifestações pessoais; b) um exercício negativo, que se revela pelo direito ao recolhimento, a se manter em segredo, de não revelar as informações pessoais (NAVES; SÁ, 2017, p. 95).

A proteção jurídica da intimidade e da privacidade deve abranger todo ser humano, seja ele nascido ou por nascer; podendo, principalmente no meio ambiente digital, se prolongar para além da morte.

O que se percebe é um ofuscamento das antigas fronteiras da privacidade, que agora se mostram mais fluidas e heterogêneas. Caminha-se entre a ampliação do acesso à informação e a necessidade de restringir o acesso a alguns dados.

Com isso, pode-se questionar qual seria a natureza jurídica formalizadora do direito à privacidade. O Direito Civil tradicional o classifica como direito subjetivo formalizado por meio de *regras jurídicas* (DE CUPIS, 2004).

A Teoria do Direito contemporânea diferencia regras e princípios jurídicos, especificados dentro do gênero “norma jurídica”, e, ainda que conviva com controvérsias, tem fixado que o princípio jurídico normalmente é uma norma mais geral, com caracteres mais abertos, que serão complementados pelo caso concreto. Assim,

O princípio apresenta sempre um agir diferente de outras normas jurídicas, uma vez que inexistem as condições prévias de aplicação. Por isso, autores como Klaus Günther entendem que princípios são normas cujas condições de aplicação não são pré-determinadas.

Günther afirma que a determinação *a posteriori* do conteúdo deve-se ao princípio possuir uma razão *on balance*, em que comparativamente avalia-se no caso concreto todas as suas condições e limites de aplicação. Sua fluidez característica leva a fundamentação a uma necessidade básica na aplicação jurídica (NAVES; REIS, 2019, p. 156-157).

Com as feições impostas pelo *Big Data*, a privacidade parece ganhar um caráter mais fluido, sem condições de aplicação predeterminadas pelo legislador, o que a encaminha para a natureza principiológica, densificada pelas condições fáticas e jurídicas para sua determinação de conteúdo.

A fim de se tutelar a privacidade quanto aos dados pessoais, em 2018, o Brasil aprovou a Lei n. 13.709 – Lei Geral de Proteção de Dados Pessoais (LGPD).

## 6 LEI GERAL DE PROTEÇÃO DE DADOS PESSOAIS

A LGPD é relativamente extensa, com dez capítulos e sessenta e cinco artigos, e objetiva regular “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 pessoa natural” (art. 1º).

Os dados pessoais são informações que identificam e caracterizam a pessoa natural e suas manifestações, por isso, incluem-se entre os direitos da personalidade, abrangendo a imagem-retrato, a imagem-atributo, a imagem-voz e a manifestação do pensamento.

*A imagem-retrato* refere-se às características fisionômicas do titular, à representação de uma pessoa pelo seu aspecto visual, enfim, é o seu pôster, a sua fotografia, encarada tanto no seu aspecto estático – uma pintura –, quanto no dinâmico – um filme –, conforme proteção dedicada pelo art. 5º, inciso X, da Constituição da República. Noutro quadrante, a *imagem-atributo* é o consectário natural da vida em sociedade, consistindo no conjunto de características peculiares da apresentação e da identificação social de uma pessoa. Diz respeito, assim, aos seus qualificativos sociais, aos seus comportamentos reiterados que permitem identificá-la. Não se confunde com a imagem exterior, cuidando, na verdade, de seu retrato moral. Já a *imagem-voz* concerne à identificação de uma pessoa através de seu timbre sonoro. Aliás, sem dúvida, a personalidade de alguém não se evidencia menos na voz que nas características fisionômicas (FARIAS; NETTO, 2017, p. 409-410).

A Lei estabeleceu o gênero “dado pessoal” e, nele, destacou a espécie “dados sensíveis”, que se relacionam a aspectos mais privados e íntimos do indivíduo, como as informações “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” (art. 5º, II).

Os dados sensíveis foram apartados na legislação por exigirem maior segurança e confidencialidade em sua utilização, posto que espelham opiniões e convicções do titular, bem como seu estado de saúde física e psíquica e sua participação política.

Com a Lei, espera-se que em pouco tempo todos os dados médicos estejam disponíveis em sistema informatizado e compartilhado entre algumas entidades de assistência à saúde.

Quanto aos dados genéticos, em razão de sua repercussão para além do titular, Reis e Oliveira (2019) explicam os princípios bioéticos, traçados

no *Human Genome Editing: science, ethics, and governance*<sup>7</sup>, e que devem nortear o uso da informações sobre o genoma, sobretudo a promoção do bem-estar, transparência, devido cuidado, ciência responsável, respeito pelas pessoas, equidade e cooperação transnacional. Alguns desses princípios bioéticos foram incorporados na LGPD, especialmente as condicionantes teleológicas de uso dos dados, o consentimento e a responsabilidade individual e coletiva.

De todo modo, como já salientado, o tratamento de dados pessoais exige o consentimento do titular ou de seu responsável legal, com referência específica do uso a que se destina. No entanto, a LGPD dispensa o consentimento em algumas hipóteses. E aí reside o problema, pois as hipóteses são muito abrangentes, como é o caso da alínea e, que estabelece a desnecessidade de consentimento quando se destinar a “proteção da vida ou da incolumidade física do titular ou de terceiro”.

Então, de maneira ampla, a proteção de terceiro pode justificar a quebra da privacidade dos dados de saúde. Mas isso não abriria as portas para uma perigosa relativização da privacidade? O que se configura como “proteção de terceiro” quanto à vida e a incolumidade física? Esse preceito permite a revelação de segredo médico?

Não há respostas a essas perguntas, mas somente a reiteração da fluidez da privacidade e da ambivalência entre o direito de saber de alguns e o direito de não querer que os outros saibam de outros.

Embora a saúde seja individual, suas repercussões podem ser geracionais, se se tratar de dados transmitidos hereditariamente.

Por outro lado, a justificativa do interesse público pode se mostrar falaciosa e autoritária. Falar da prevalência de um interesse público sobre o interesse particular é considerar a existência de um interesse universalizável, o que se torna, no mínimo, muito difícil nas sociedades plurais contemporâneas.

Se, por outro lado, se escolher o argumento do interesse da maioria, correm-se dois grandes perigos: o perigo da tirania da maioria, que não respeita nem os direitos fundamentais para sobressair-se aos demais; o perigo da consideração de que o todo não é formado por partes.

Ora, se há esse interesse público, como valor predominante na sociedade, ele não é um interesse estatal, mas um conjunto de interesses privados, que se formam na inconstância histórica dos argumentos.

<sup>7</sup> Relatório divulgado em 2017 e redigido pelas *National Academies of Sciences, Engineering, and Medicine* dos Estados Unidos.

Daniel Sarmiento defende a impropriedade da “supremacia do interesse público”:

Portanto, o quadro que se delinea diante dos olhos é muito mais o de convergência entre interesses públicos e particulares do que o de colisão. Tal situação, repita-se, não constitui a exceção, mas a regra. Na imensa maioria dos casos, a coletividade se beneficia com a efetiva proteção dos interesses dos seus membros. Até porque, o interesse público, na verdade, é composto pelos interesses particulares dos membros da sociedade, razão pela qual se torna em regra impossível dissociar os interesses públicos dos privados (SARMENTO, 2007, p. 83-84).

Outro ponto relevante da LGPD é o compartilhamento de dados sensíveis, que, sob certas circunstâncias, é permitido.

Em uma leitura rápida, parece não ser permitido o compartilhamento de dados pessoais sensíveis referentes à saúde com objetivo de obter vantagem econômica. No entanto, o § 4º do art. 11 abre tantas exceções a essa proibição, que ela mais aparenta permitir do que o contrário:

§ 4º É vedada a comunicação ou o uso compartilhado entre controladores de dados pessoais sensíveis referentes à saúde com objetivo de obter vantagem econômica, **exceto nas hipóteses relativas a prestação de serviços de saúde, de assistência farmacêutica e de assistência à saúde**, desde que observado o § 5º deste artigo, incluídos os serviços auxiliares de diagnose e terapia, em benefício dos interesses dos titulares de dados, e para permitir:

I – a portabilidade de dados quando solicitada pelo titular;

II – as transações financeiras e administrativas resultantes do uso e da prestação dos serviços de que trata este parágrafo.

Tal compartilhamento, todavia, é vedado para as operadoras de planos de saúde e outras atividades econômicas que não estejam relacionadas a serviços de saúde (§ 5º do art. 11 da LGPD).

## CONCLUSÃO

No meio ambiente digital, os dados são constituintes da própria identidade, transfigurada em *bytes* para possibilitar o armazenamento em uma única linguagem.

O *Big Data* impede que seres humanos tenham, diretamente, a possibilidade de mapear e cruzar dados, o que leva a humanidade para o inevitável caminho da inteligência artificial.

Por essa razão, é necessário que o Direito estabeleça regulações e procedimentos para o armazenamento, o controle, a divulgação e o

compartilhamento de dados, especialmente os dados de pessoas naturais e jurídicas.

Nesse intuito, algumas normas vêm sendo formuladas, como a LGPD brasileira, com o objetivo de preservar a liberdade de expressão e de comunicação, mas com garantias aos direitos à privacidade, à honra, à imagem e à livre concorrência.

O direito à privacidade, como manifestação do espaço individual a que se escolhe compartilhar restritivamente, ganha novas feições sob a influência do meio ambiente digital, pois, ao mesmo tempo em que ganha novas possibilidades de manifestação, também abre espaço a novas violações. Em se tratando da rede mundial de computadores, dados que foram divulgados livremente não retornam mais ao círculo privado, já que o controle dos dados pode ser tratado como propriedade, a que se pretende completo domínio.

O assunto ainda pode ganhar conotações mais dramáticas diante de dados sensíveis, que revelam informações geracionais, médicas, convicções religiosas e opiniões políticas. Sobre tais dados, embora prevaleça a regra da confidencialidade – logo sua divulgação exige o consentimento do titular –, haverá situações em que sua revelação poderá se tornar útil à saúde de terceiro e as fronteiras do direito terão de se projetar ou se retrair diante das circunstâncias fáticas.

Com isso, a clássica configuração das normas atinentes à privacidade como regras jurídicas parecem cair por terra, dando lugar à fluidez principiológica – princípio da privacidade –, sem a possibilidade de o legislador descrever, como em uma regra-tipo ou suporte fático, as suas condições de aplicação.

## REFERÊNCIAS

BRASIL. *Código Civil*. Lei n. 10.406. 10 jan. 2002. Institui o Código Civil. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/leis/2002/L10406compilada.htm](http://www.planalto.gov.br/ccivil_03/leis/2002/L10406compilada.htm). Acesso em: 1 mar. 2020.

BRASIL. Constituição (1988). *Constituição da República Federativa do Brasil*. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/constituicao/constituicaocompilado.htm](http://www.planalto.gov.br/ccivil_03/constituicao/constituicaocompilado.htm). Acesso em: 1 mar. 2020.

BRASIL. Lei n. 13.709, de 14 de agosto de 2018. *Lei Geral de Proteção de Dados Pessoais* (LGPD). Disponível em: [http://www.planalto.gov.br/ccivil\\_03/\\_ato2015-2018/2018/lei/L13709.htm](http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm). Acesso em: 1 mar. 2020.

DE CUPIS, A. *Os direitos da personalidade*. Campinas: Romana Jurídica, 2004.

DUHIGG, C. *O poder do hábito: por que fazemos o que fazemos na vida e nos negócios*. Rio de Janeiro: Objetiva, 2012.

FARIAS, C. C.; NETTO, F. B.; ROSENVALD, N. *Novo tratado de responsabilidade civil*. 2. ed. São Paulo: Saraiva, 2017.

FIORILLO, C. A. P. *Curso de direito ambiental brasileiro*. 13. ed. rev., atual. e ampl. São Paulo: Saraiva, 2012.

FLORIDI, L.; TADDEO, M. What is data ethics? *Philosophical Transactions of the Royal Society a Mathematical, Physical and Engineering Sciences*, v. 374, p. 1-5, dec. 2016. Disponível em: <https://royalsocietypublishing.org/doi/pdf/10.1098/rsta.2016.0360>. Acesso em: 30 jan. 2020.

GINSBERG, J. et al. Detecting influenza epidemics using search engine query data. *Nature*, v. 457, p. 1012-1014, feb. 2009. Disponível em: <https://www.nature.com/articles/nature07634>. Acesso em: 30 jan. 2020.

GOLDER, S. A; MACY, M. W. Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science*, v. 333, p. 1878-1881, sep. 2011. Disponível em: <https://science.sciencemag.org/content/333/6051/1878.full>. Acesso em: 31 jan. 2020.

LANEY, D. *3D data management: controlling data volume, velocity, and variety*. Disponível em: <https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>. Acesso em: 31 jan. 2020.

MAYER-SCHONBERGER, V.; CUKIER, K. *Big data: the essential guide to work, life and learning in the age of insight*. New York: Houghton Mifflin Harcourt, 2013.

NAVES, B. T. O.; REIS, É. V. B. *Bioética ambiental: premissas para o diálogo entre a ética, a bioética, o biodireito e o Direito Ambiental*. 2. ed., rev. e aum. Rio de Janeiro: Lumen Juris, 2019.

NAVES, B. T. O.; SÁ, M. F. F. *Direitos da personalidade*. Belo Horizonte: Arraes, 2017.

REIS, É. V. B.; OLIVEIRA, B. T. CRISPR-CAS9, biossegurança e bioética: uma análise jusfilosófica-ambiental da engenharia genética. *Veredas do*

*Direito*, Belo Horizonte, v. 16, n. 34, p. 123-152, jan./abr. 2019.

SAMUEL, A. L. Some studies in machine learning using the game of checkers. *IBM Journal*, v. 3, p. 210-229, jul. 1959. Disponível em: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5392560>. Acesso em: 1 fev. 2020.

SARMENTO, D. Interesses públicos vs. Interesses privados na perspectiva da teoria e da filosofia constitucional. In: SARMENTO, D. (Org.). *Interesses públicos versus interesses privados: desconstruindo o princípio da supremacia do interesse público*. Rio de Janeiro: Lumen Juris, 2007. p. 23-116.

TAYLOR, C. *As fontes do self: a construção da identidade moderna*. São Paulo: Loyola, 1997.

TAULLI, T. *Introdução à inteligência artificial: uma abordagem não técnica*. São Paulo: Novatec, 2020.

Artigo recebido em: 04/03/2020.

Artigo aceito em: 06/05/2020.

### **Como citar este artigo (ABNT):**

REIS, E. V. B.; NAVES, B. T. O. O meio ambiente digital e o direito à privacidade diante do Big Data. *Veredas do Direito*, Belo Horizonte, v. 17, n. 37, p. 145-167, jan.-abr. 2020. Disponível em: <http://revista.domhelder.edu.br/index.php/veredas/article/view/1795>. Acesso em: dia mês. ano.