

1 INTRODUÇÃO

A revolução nos meios de produção, aliada ao progresso da tecnologia, desencadearam a formação de uma nova sociedade que passou a se organizar de maneira distinta, por exemplo, com a difusão da internet, as relações sociais tomaram outros espaços, agora virtuais, e a sociedade passou a conviver em rede¹.

Com a investida tecnológica, surgem, a todo instante, redes sociais que exigem interação. Pessoas ficam cada vez mais reféns de seus *smartphones* e seus aplicativos, que revelam o perfil de consumidor graças ao uso constante do cartão de crédito seja em estabelecimentos físicos ou em lojas virtuais. Faz-se imprescindível manter-se conectado.

A vida passa a concentrar-se na velocidade intransponível e incessante em uma sociedade na qual ficar fora dos bancos de dados significa não existir como sujeito, tudo isso tornado possível mediante a relação entre tecnologia e economia.

A internet acabou por conferir certa notoriedade a cada indivíduo, que antes dela muitas vezes era invisível, e tornou “[...] tangível e irrefutável o ser e o estar no mundo” (BAUMAN, 2014, p. 88). Isso explica a facilidade que cada sujeito tem em dispor sobre suas informações pessoais, seja para sentir-se notado, seja para adquirir bens e serviços. O fator econômico, que se sobressai na sociedade capitalista, é uma das molas propulsoras que alimentam os bancos de dados automatizados e que fazem com que grandes corporações estejam cada vez mais interessadas em se apropriar das novas tecnologias como as análises de *big data*, o que, em contrapartida, demanda pensar na proteção jurídica dos dados pessoais.

Adepta das novas tecnologias, a sociedade acabou por concebê-las como instrumentos ora bons, ora ruins, mas sempre em sua neutralidade, e eficácia e nunca como arma que domina e destrói, o que se denota, até então, uma relação de dependência. Isto porque o indivíduo, por meio das tecnologias, desenvolvem mecanismos que buscam facilitar sua vida, significando dizer que as inovações tecnológicas têm valor para a sociedade, seja ele econômico, social ou científico (ECO, 2001).

Nesse contexto, o presente artigo tem como finalidade examinar o *big data* no contexto da sociedade de controle, através do pensamento do filósofo Gilles Deleuze, como

¹ O termo “rede” aqui suscitado refere-se às redes sociotécnicas que o filósofo Rogério Costa, professor da PUC-SP, descreve como: “[...] água, transportes, comércio, telecomunicação, telefonia, comunicação, TV, jornal, computação, *web*, portáteis. Estamos dentro de muitas redes simultaneamente e permanentemente. [...] Na cidade digital, em casa ou no trabalho, pelo fato de essas redes estarem interconectadas, podemos acessar múltiplos serviços sem a necessidade de nos deslocarmos. Temos entrega de produtos, pagamentos tipo *homebanking*, serviços públicos, trabalho e muitas outras coisas possíveis pelo fato de que a cidade está digitalizada” (COSTA, Rogério da. Sociedade de controle. **São Paulo Perspec.** São Paulo, v. 18, n. 1, p. 161-167, mar. 2004. Disponível em <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-88392004000100019&lng=pt&nrm=iso>. Acesso em 20 ago. 2018, p. 166, excluímos).

desdobramento da sociedade disciplinar objeto dos estudos de Michel Foucault, para, assim, situar os bancos de dados automatizados como projeto arquitetônico de vigilância pós-moderno. Após, traz-se uma definição mais técnica do *big data* correlacionando com as suas principais características que auxiliam na sua compreensão. Por fim, abordam-se os valores científico, econômico e social do *big data*, refletidos nos bens e serviços tornados possíveis por essa nova tecnologia, articulando com os riscos à privacidade intrínsecos ao processo de coleta, tratamento e uso de dados pessoais.

A fim de alcançar os resultados pretendidos, utilizou-se o método dedutivo, visto que o trabalho tem a finalidade de explicar o conteúdo das ideias expostas, partindo de uma concepção geral para uma específica, de um viés filosófico para viés técnico na configuração do *big data*, com atenção aos riscos para a privacidade, por meio do tratamento de dados pessoais tornado possível a partir dessa nova tecnologia. A abordagem metodológica utilizada foi a qualitativa que permite o emprego de “[...] diferentes concepções filosóficas; estratégias de investigação; e métodos de coleta, análise e interpretação dos dados” (CRESWELL, 2006, p. 206).

2 DA SOCIEDADE DISCIPLINAR PARA A SOCIEDADE DE CONTROLE: A VIGILÂNCIA CONSTANTE

De antemão, cabe destacar aqui que não se percebe uma recusa da tecnologia, mas sim a intenção de, a partir do exame dos usos e seus efeitos, situar como suas inovações podem conviver, ou melhor, adaptar-se à máxima eficiência dos bens e serviços que oferecem, sem ultrapassar as barreiras que protegem a privacidade do indivíduo na sua condição de usuário e membro-usuário da nova sociedade, que pode ser exemplificada a partir da passagem do mundo analógico para o digital tratado por Paula Sibilia que, para exemplificar o impulso massificante e individualizante² da sociedade industrial, cita a “carteira de identidade”, considerando que,

Tal documento faz referência a um Estado-nação, contém um número que localiza o indivíduo dentro da massa, uma foto, uma assinatura e uma impressão do dedo polegar. Todos dados analógicos. De outro lado, o sujeito da sociedade contemporânea detém cada vez mais cartões de crédito e senhas de acesso: dispositivos digitais. De maneira crescente, a identificação

² Paula Sibilia explica que “[...] na sociedade contemporânea, tanto a noção de massa quanto a de indivíduo estão perdendo força. No lugar dessas figuras, outras emergem. O papel do consumidor, por exemplo, assume uma relevância cada vez maior. Em vez de integrar uma massa - como os cidadãos dos Estados-nação da era industrial - ele faz parte de diversas amostras, nichos de mercado, segmentos de público, *targets* e bancos de dados.” (SIBILIA, Paula. **O homem pós-orgânico**: o corpo, subjetividade e tecnologias digitais. Rio de Janeiro: Relume Dumará, 2002, p. 34)

do consumidor para pelo perfil: uma série de dados sobre a condição socioeconômica, seus hábitos e suas preferências de consumo, colhidos através de formulários de pesquisas e processados digitalmente, para serem armazenados em bancos de dados conectados em rede que serão acessados, vendidos, comprados e utilizados pelas empresas em suas estratégias de marketing (SIBILIA, 2002, p. 34).

Neste cenário das novas tecnologias, a partir da criação dos bancos de dados automatizados, nas quais as informações disponibilizadas em rede são colhidas e armazenadas pelas grandes empresas, na maioria das vezes sem a autorização do usuário, ou até que se admitam casos em que há expresso consentimento, ainda assim sem o conhecimento³ acerca dos usos secundários que serão feitos dos dados disponibilizados. É nesse cenário que surge o *big data*⁴. Essa nova tecnologia consiste em análises que correlacionam os dados armazenados dos indivíduos e revelam informações pessoais em uma quantidade e velocidade jamais vista, não há mais tempo para descanso e tranquilidade, mas apenas a constante precaução ao mover-se na era em que os dados, as informações adquirem valor supremo.

Nesse contexto, percebe-se que os bancos de dados automatizados aliados às análises de *big data* revelam uma forma de controle e vigilância dos indivíduos, pelos seus gestores, sobre toda a sociedade, a qual, diariamente, vê captadas informações pessoais a partir da simples conexão à internet. Assim, para melhor situar a discussão a cerca do que constitui o *big data* e a necessidade de uma efetiva proteção jurídica aos dados pessoais no cenário das novas tecnologias, cabe examinar a transição das sociedades disciplinares para as sociedades de controle. Esta última posicionada como receptora da nova tecnologia, sendo os grandes bancos de dados automatizados como o seu novo Panóptico, em uma metáfora à

³ Uma pesquisa publicada na *Harvard Business Review* Brasil em fevereiro de 2016 aponta que as pessoas, apesar de terem pleno conhecimento de que suas informações pessoais são coletadas pelas empresas com que mantêm relação, surpreendem-se com a falta de informação quanto aos tipos específicos de dados que fornecem; a pesquisa aponta que, das pessoas entrevistadas, 27% percebem que compartilham a sua lista de amigos em redes sociais, 25% que compartilham a sua localização, 23% suas buscas na *web*, 18% seus históricos de comunicação como “chat logs”, 17% seus endereços de IP e 14% percebem que compartilham seu histórico de navegação na *web*. (MOREY, Timothy; FORBATH, Theodore; SCHOOP, Alisson, Dados dos consumidores: modelos de transparência e confiança. **Harvard Business Review Brasil**, São Paulo, fevereiro 2016, p. 47)

⁴ O termo *big data* foi cunhado pelas ciências como a astronomia e a genômica, que vivenciaram uma explosão nos anos 2000; atualmente tem migrado para todos os campos do conhecimento humano e ainda não se tem uma definição rigorosa para o termo, considerando as suas características Viktor Mayer-Schonberger e Kenneth Cukier consideram o *big data* como um conjunto de tendências tecnológicas que consiste em uma abordagem que aplica à matemática uma enorme quantidade de dados a fim de prever probabilidades, permitindo analisar muito mais dados, sem devoção ao rigor da exatidão, buscando descobrir novos padrões e correlações nos dados que propiciam novas e valiosas ideias. Outros autores como Zikopoulos et al. (2012) dizem que *big data* se caracteriza por quatro aspectos: volume, velocidade, variedade e veracidade. (MAYER-SCHONBERGER, Viktor; CUKIER, Kenneth. **Big data**: como extrair volume, variedade, velocidade e valor da avalanche de informação cotidiana. Tradução Paulo Polzonoff Junior. Rio de Janeiro: Elsevier, 2013; ZIKOPOULOS, P; DE ROOS, D; PARASURAMAN, K; DEUTSCH, T; GILES, J; CORRIGAN, D. **Harness the power of Big Data**-The IBM Big Data Platform. Emeryville: McGraw-Hill Osborne Media, 2012).

construção de Jeremy Bertham utilizado por Michel Foucault para descrever as relações de poder nas sociedades disciplinares.

Para compreender a inserção e atuação das novas tecnologias, especificamente do *big data*, na nova sociedade incitada pelos impulsos tecnocráticos que seguiram o pós-guerra, é indispensável se considerem dois filósofos que acreditavam na evolução do pensamento por meio de crises: Gilles Deleuze e Michel Foucault. O primeiro interlocutor do pensamento do segundo, dando um prosseguimento às ideias do outro. Enquanto Foucault debruçou-se sobre a chamada sociedade disciplinar, Deleuze introduziu o que chama de sociedade de controle⁵, marcando a transição da primeira para a segunda, quando afirma que “[...] são as *sociedades de controle* que estão substituindo as sociedades disciplinares” (DELEUZE, 1992, p. 224).

Não é objetivo de Deleuze retratar o futuro com ficção, mas, diagnosticar o presente lançando uma nova imagem do pensamento sobre a contemporaneidade. Eis que, na sua leitura, surgem as sociedades de controle como desdobramento das sociedades de soberania e disciplinares demonstradas do século XVII ao início do XX, por Foucault.

Para Braga e Vlach (2004, n.p) “[...] devemos entender a sociedade de controle como aquela (que se desenvolve nos limites da modernidade e se abre para a pós-modernidade)”. Nesse interim, Silvana Tótorá (2006, p. 239) afirma que “[...] no capitalismo pós-industrial, os próprios homens são peças constitutivas da máquina e não seus meros usuários; são partes de uma engrenagem de circulação de informação”.

No seu texto *Post-Scriptum sobre as Sociedades de Controle*, Gilles Deleuze, aparentemente, faz com que o leitor imagine uma ficção científica. No entanto não é essa sua intenção, por isso esclarece não haver necessidade de ficção científica “[...] para se conceber um mecanismo de controle que dê, a cada instante, a posição de um elemento em espaço aberto, animal sem reserva, homem numa empresa (coleira eletrônica)” (DELEUZE, 1992, p. 228-229).

Nas chamadas sociedades disciplinares, imperava o espaço fechado da disciplina dos corpos, a fim de torná-los dóceis. Por esse motivo “[...] forma-se então uma política das coerções que são um trabalho sobre o corpo, uma manipulação calculada de seus elementos, de seus gestos, de seus comportamentos” (FOUCAULT, 2010, p. 133). A disciplina está associada aos espaços fechados aos quais o indivíduo está submetido e deles migra constantemente, “[...] primeiro a família, depois a escola [...], depois a caserna [...], depois a

⁵ Deleuze explica, ao referir-se à sociedade de controle, que o termo “controle” foi proposto por Burroughs para designar o novo monstro e que Foucault reconhece como o futuro próximo; menciona ainda outro autor que também analisa as formas de controle que funcionam como substitutas das antigas disciplinas, referenciando Paul Virilio (DELEUZE, G. *Post-scriptum sobre as sociedades de controle*. In: _____. **Conversações (1972-1990)**. São Paulo: Ed. 34, 1992, p. 224).

fábrica, de vez em quando o hospital, eventualmente a prisão, que é o meio de confinamento por excelência” (DELEUZE, 1992, p. 223; suprimimos).

Denota-se que na sociedade disciplinar a vigilância adquire um aspecto determinante para o encarceramento que isola o olhar. Abre-se espaço para a impossibilidade de se escapar ao olhar do outro, em razão das aberturas que vazam os interiores e colocam o indivíduo em uma situação de constante observação (CHEVITARESE; PEDRO, 2002, pp. 129-162).

De uma sociedade para outra, o homem passa de confinado para endividado (DELEUZE, 1992). Endividado, porque nas sociedades de controle os homens estão submetidos a um controle contínuo, sem delimitação de início, meio e fim, num processo modificado a todo instante.

As diferenças entre sociedades de disciplina e sociedades de controle são estabelecidas por Deleuze como sendo característico da primeira o constante recomeço, fosse ele da escola à caserna, da caserna à fábrica, enquanto que na segunda, nunca se termina nada. Menciona que “[...] as sociedades disciplinares têm dois polos: a assinatura que indica o *indivíduo*, e o número de matrícula que indica sua posição numa *massa*” (DELEUZE, 1992, p. 226; destaque no original), diferente do que ocorre nas sociedades de controle, nas quais “[...] a linguagem numérica do controle é feita de cifras, que marcam o acesso à informação, ou a rejeição” (DELEUZE, 1992, p. 226).

Nas sociedades de controle, o indivíduo tem suas informações registradas a partir de uma cifra que, como uma senha, fornece acesso a todas as informações. Isto é possível em detrimento da geração de informação cibernética e computadorizada visto que, “[...] as máquinas da informática e os computadores não são apenas evoluções tecnológicas, mas operam uma mutação no capitalismo” (BRAGA; VLACH, op. cit., 2004, n.p.). As tecnologias detectam a posição de cada um, lícita ou ilícitamente, e operam uma modulação universal, através do controle contínuo e da comunicação instantânea (DELEUZE, 1992).

Considerando os elementos e configuração da sociedade de controle, percebe-se que esta tem “[...] como valores máximos de mercadoria o prestígio, a informação, o conhecimento e como dispersão máxima a força de trabalho” (NEVES, 1997, p. 87). É na sociedade de controle que “os indivíduos tornaram-se *‘dividuais’*”, divisíveis, e as massas tornaram-se amostras, dados, mercados ou *bancos*” (DELEUZE, op. cit., 1992, p. 226; destaque no original).

Acredita-se que a nova caracterização do indivíduo deve-se principalmente às máquinas que operam a sociedade de controle, notadamente as máquinas de informática e computadores, nas quais operam e imperam a comunicação virtual e o controle de dados

peçoais como forma de vigilância. Quanto às sociedades que se sucederam na história, atribuem-se a elas diferentes máquinas que, respectivamente, as fizeram operar, como exemplo, para as sociedades de soberania estavam as máquinas simples, alavancas, relógios; para as sociedades disciplinares as máquinas energéticas (DELEUZE, 1992).

A partir do endividamento do indivíduo, que se encontra submetido a um controle contínuo, no qual imperam os computadores e a informática, tornando os sujeitos seres divisíveis, torna possível perceber a permanência da “vigilância”, a mesma retratada por Foucault em sua obra *Vigiar e Punir*, com a figura do Panóptico⁶, de enorme valor ao poder disciplinar, cujo efeito mais importante era “[...] induzir no detento um estado consciente e permanente de visibilidade que assegura o funcionamento automático do poder” (FOUCAULT, 2010, p. 191).

Para estabelecer a relação entre o Panóptico de Jeremy Bentham, utilizado por Foucault, com os bancos de dados, Bauman (1999) recorre a um ensaio de Mark Poster que trata os bancos de dados eletrônicos como uma versão ciberespacial atualizada do Panóptico, nos quais os indivíduos seriam fisgados dentro das redes, dos bancos de dados, sem poder apresentar qualquer resistência. Estaríamos, pois, diante do “superpanóptico”, mas com uma diferença “[...] os vigiados, fornecendo os dados a armazenar, são fatores primordiais – e voluntários – da vigilância” (BAUMAN, op. cit., 1999, p. 58; destaque no original).

Ao traçar a diferença entre o Panóptico e os bancos de dados, Bauman (1999) afirma que o primeiro tem como função primordial evitar que qualquer pessoa escape do espaço estritamente vigiado, enquanto que o segundo tem como principal função garantir que nenhum indivíduo entre nos bancos de dados sob falsas alegações ou sem credibilidade. O autor associa o fato de que os bancos de dados representam uma possibilidade de conferir ao indivíduo certa credibilidade, haja vista que são criados e geridos por empresas que têm interesses em traçar o perfil do indivíduo como consumidor.

⁶ Foucault (2010) explica a arquitetura do panóptico que atribui a Bentham; segundo o autor, consiste “[...] na periferia uma construção em anel; no centro, uma torre: esta é vazada de largas janelas que se abrem sobre a face interna do anel; a construção periférica é dividida em celas, cada uma atravessando toda a espessura da construção; elas têm duas janelas, uma para o interior, correspondendo às janelas da torre; outra, que dá para o exterior, permite que a luz atravesse a cela de lado a lado. Basta então colocar um vigia na torre central, e em cada cela trancar um louco, um doente, um condenado, um operário ou um escolar. Pelo efeito da contraluz, pode-se perceber da torre, recortando-se exatamente sobre a claridade, as pequenas silhuetas cativas nas celas da periferia. Tantas jaulas, tantos pequenos teatros, em que cada ator está sozinho, perfeitamente individualizado e constantemente visível. O dispositivo panóptico organiza unidades espaciais que permitem ver sem parar e reconhecer imediatamente. Em suma, o princípio da masmorra é invertido; ou antes, de suas três funções – trancar, privar de luz e esconder – só se conserva a primeira e suprimem-se as outras duas. A plena luz e olhar de um vigia captam melhor que a sombra, que finalmente protegia. A visibilidade é uma armadilha” (FOUCAULT, op. cit.2010, p. 190).

Os bancos de dados automatizados, aliados às análises de *big data*, possibilitam que a prisão dos indivíduos dispense um lugar fixo e fechado, pois permitem que, em qualquer lugar, seja aberto ou fechado, todos estejamos sob controle pela vigilância daqueles que rastreiam os passos, correlacionando dados. Um dos maiores exemplos que confirmam esta assertiva é o acesso às informações do *Global Positioning System* (GPS), as quais descrevem as rotas e lugares mais frequentados pelas pessoas, a ponto de identificar o trabalho, a residência, a escola dos filhos, entre outros lugares.

Para demonstrar como os bancos de dados automatizados representam o novo Panóptico na sociedade de controle, basta observar o projeto americano *Total Terrorism Information Awareness* (TIA), cujo objetivo é colher o máximo de informações pessoais acessando registros financeiros, médicos, de comunicação, viagens, entre outros, associando todas essas informações e criando um grande banco de dados, virtual e centralizado. Através desse banco de dados, o projeto poderá rastrear os indivíduos para detectar suas atividades potenciais, com isso, percebe-se que “[...] o reconhecimento de padrões está diretamente ligado à mudança nos métodos de controle das ações individuais” (COSTA, 2004, p. 165).

O projeto americano TIA é um exemplo claro de inovação, tornado possível devido à tecnologia do *big data*. Com isso, é imperioso observar que há um valor intrínseco em cada banco de dados, muitas vezes oculto, pois, como expõem Viktor Mayer-Schonberger e Kenneth Cukier (2013), desde o século XX os dados são o petróleo da economia da informação. Isto justifica a corrida para descobrir e captar esse valor intrínseco, abrindo espaço para violações à privacidade dos indivíduos, quando nestes bancos de dados estão presentes dados pessoais. No contexto do *big data*, as máquinas acabam por possibilitar a criação de modelos estatísticos onde se encaixam aos dados, otimizam-nos e os preveem (DAVENPORT, 2014).

Ao tratar do modelo Panóptico do poder moderno concebido por Michel Foucault, Zygmunt Bauman (1999) esclarece que esse projeto arquitetônico de vigilância tinha como propósito manipular e rearrumar a transparência do espaço como relação social, e até como relação de poder. Isto porque os internos estavam em total situação de visibilidade e os supervisores em toda invisibilidade, fazendo com que os primeiros se comportassem como se estivessem em constante vigilância.

Retratando a vigilância como disfarçada de serviço da segurança, David Lyon, em diálogo travado com Bauman, problematiza a questão do panóptico pós-moderno, a partir do contexto que insere:

Os olhos eletrônicos sempre abertos nas ruas, a coleta de dados abrangente, os fluxos de informações pessoais com sua pressão cada vez mais alta são

vistos como reações racionais aos riscos da vida. Precisamos desesperadamente de vozes que perguntem: “Por quê?” “Para quê?” “Você tem alguma ideia das consequências disso tudo?” Eu fico atento, esperando ouvir alguém dizer: “Haveria outras maneiras de conceber o que há de errado com o mundo, e como seus males poderiam ser abordados?” (BAUMAN, 2014, p. 79-80).

As provocações de David Lyon são mais que pertinentes, tornaram-se necessárias. Ainda que não se tenham respostas prontas - e essa sequer seria a melhor saída, face à dinamicidade dos fenômenos tecnológicos - o “por quê?”, “para quê?” e “quais as consequências?” da utilização das novas tecnologias, como o *big data*, é imprescindível na discussão sobre uma efetiva proteção jurídica aos dados pessoais, a fim de que seja possível refletir acerca de como se tem constituído a nova vigilância na sociedade de controle, a partir das tecnologias da informação e, analisando os instrumentos jurídicos disponíveis, seja possível averiguar como podem ser garantidas a preservação da liberdade do indivíduo e sua autonomia em relação aos seus dados pessoais, sem obstar o avanço tecnológico, na conjuntura da sociedade da informação que segundo Manuel Castells (2006, p. 268), organiza

[...] seu sistema produtivo em torno de princípios de maximização da produtividade baseada em conhecimentos, por intermédio do desenvolvimento e da difusão de tecnologias da informação e pelo atendimento dos pré-requisitos para sua utilização (principalmente recursos humanos e infra-estrutura de comunicações).

Ante o exposto, resta concluir que Foucault ilumina as conexões entre o Panóptico e a modernidade, em que as figuras da punição e reforma das práticas disciplinares no discurso pós-moderno, evidenciado por Bauman, tomam outra proporção, abrindo espaço para a vigilância como punição, e a prática disciplinar rompe as barreiras dos espaços confinados ao que, anteriormente, estava restrita a figura do panóptico, para adotar uma nova roupagem: a vigilância eletrônica, na qual os detalhes pessoais da vida dos indivíduos passam a circular fora no seu controle para dentro de bancos de dados remotos e automatizados que representam uma espécie de versão “ciberespacial” do Panóptico, que tem sua concretização arquitetônica, a exemplo, a partir dos *data centers*⁷ do *Google*. Assim, vivemos na sociedade de controle, movidos pela difusão da tecnologia que nos insere na sociedade da informação.

3 O *BIG DATA* E SUA INFLUÊNCIA NO TRATAMENTO E ARMAZENAMENTO DE DADOS PESSOAIS

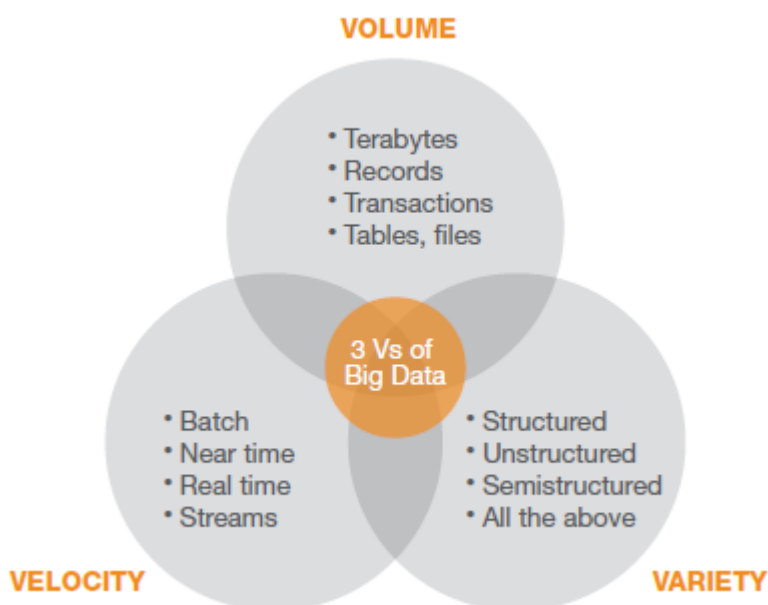
⁷ Para fazer um tour em um dos principais *data centers* do provedor *Google*, que fica em The Dalles acesse: <<http://exame.abril.com.br/tecnologia/noticias/conheca-o-data-center-do-google-que-nem-funcionarios-entram>> Acesso em 20 ago. 2018.

Os avanços dos investimentos nas tecnologias de informação, a criação da rede mundial de computadores e todas as tecnologias digitais fizeram com que as informações encontrassem maior propulsão para circular e se desenvolverem. Conseqüentemente, com esse novo cenário da sociedade da informação, novas tecnologias surgiram com o intuito de modernizar e amparar as necessidades humanas, e, uma delas repousa na elaboração de ideias e criação de bens e serviços, a partir da correlação de dados que se encontram em grandes bancos de dados, ou seja, o *big data*.

A impressionante velocidade da escalada tecnológica, desta vez, ditou uma tendência em tecnologia que está abrindo caminho para um novo método de compreensão do mundo, o *big data*. Bilhões de consumidores deixam rastros digitais ao fazerem pesquisas no *Google*, editar comentários em redes sociais, acessar vídeos no *YouTube* ou utilizar o GPS do automóvel. Esses dados estão gerando um imenso banco de dados, denominado *big data* (FEIJÓ, 2013, pp. 30-43).

Ao falar sobre *big data*, a quantidade dos dados importa, mas não é tudo. Assim explica Philip Russom (2011) ao considerar que outras características do *big data*, as quais são imprescindíveis para definir essa nova tecnologia, referindo-se a: *volume*, *velocidade* e *variedade*, os três Vs, ressaltando que cada característica possui suas ramificações, conforme se analisa na figura a seguir.

Figura 1 - Características que definem o *big data*



Fonte: RUSSOM, 2011, p. 6.

O principal atributo do *big data* é o *volume*⁸ de dados que está envolvido nas análises feitas por meio dessa tecnologia, razão pela qual a maioria das pessoas costuma definir *big data* em *terabytes* e às vezes *petabytes*. Explica Philip Russom (2011) que o *big data* não é só quantificado pelo volume dos dados, mas também por contagem de registros, transações, tabelas ou arquivos. Isto porque, muitas vezes, nem todos os dados coletados para armazenamento serão necessariamente coletados para análise. Além disso, em razão da quantificação ser crescente, é difícil quantificar, logo, a definição de *big data* não se restringe apenas o volume de dados, representando assim, uma das suas características que, aliada a outras, conseguem defini-lo com maior propriedade. Para melhor compreender esse volume de dados no qual se insere o *big data*, Philip Chen e Chun-Yang Zhang (2014, p. 3) descrevem que

Na era da informação, quase todas as grandes empresas encontram problemas com Big Data, especialmente as corporações multinacionais. Por um lado, essas empresas, em sua maioria, têm um grande número de clientes em todo o mundo. Por outro lado, há muito grande volume e velocidade de transação de dados. Por exemplo, o sistema de detecção de fraude de cartão de crédito Falcon da FICO gerencia mais de 2,1 bilhões de contas válidas em todo o mundo. Há acima de 3 bilhões de pedaços de conteúdo gerados no Facebook todos os dias. O mesmo problema acontece em todas as empresas de Internet. A lista poderia continuar, como nós testemunhamos os campos futuros da batalha dos negócios que focalizam em Big Data.

Sabendo que o *big data* foi criado para analisar grande *volume* de dados, seu grande diferencial está explicado na segunda característica que é a maior *variedade* das fontes (*web*, mídias sociais, aparelhos eletrônicos, entre outros) utilizadas para coletar todo tipo de informação, seja pessoal ou não. Com a variedade das fontes, os usuários de *big data*, as corporações, passaram a não simplesmente acumular os dados coletados, como analisá-los para extrair e direcionar bens e serviços, o *big data* trouxe para as corporações uma mudança significativa:

As poucas organizações que têm vindo a analisar estes dados agora o fazem a um nível mais complexo e sofisticado. O *big data* não é novo, mas o alavancamento analítico efetivo do *big data* é. A recente exploração dessas fontes para análise significa que os dados chamados estruturados (que anteriormente possuíam hegemonia incontestada na análise) agora são associados por dados estruturados (texto e linguagem humana) e dados semiestruturados (XML, *feeds* RSS). Há também dados que é difícil de categorizar, como os de áudio, vídeo e outros dispositivos (RUSSOM, 2011, p. 7).

⁸ “The thing about Big Data and data *volumes* is that the language has changed. Aggregation that used to be measured in petabytes (PB) is now referenced by a term that sounds as if it’s from a Star Wars movie: zettabytes (ZB). A zettabyte is a trillion gigabytes (GB), or a billion terabytes!” (ZIKOPOULOS et al., op. cit., 2012, p. 9). Disponível em: <<http://www.dotgroup.co.uk/wp-content/uploads/2014/11/Harness-the-Power-of-Big-Data-The-IBM-Big-Data-Platform.pdf>>. Acesso em 20 ago. 2018.

A *variedade* como característica do *big data* refere-se, portanto, à capacidade de capturar todos os dados necessários para a tomada de decisões, independente de onde esteja, incluindo assim até opiniões no *Facebook*, imagens e outros, dados estruturados, semiestruturados ou não estruturados⁹, que Zikopoulos et al. (2012, p. 10) esclarecem que “[...] todos os dados têm alguma estrutura, quando nos referimos a dados não estruturados, estamos nos referindo aos subcomponentes que não têm estrutura, como o texto de forma livre em um campo de comentários.”¹⁰

Aliado ao *volume* dos dados e à *variedade* das fontes de coleta, a *velocidade* com que são gerados, entregues e processados os dados é o terceiro atributo do *big data*. A *velocidade* torna possível a análise de todo o fluxo de dados coletados num fluxo contínuo onde há necessidade de dar sentido aos dados e, ao mesmo tempo, tomar decisões em tempo real (RUSSOM, 2011). Essa característica é também definida como “[...] a taxa na qual os dados chegam à empresa e são processados ou bem compreendidos [...]” (ZIKOPOULOS et al., op. cit 2012, p. 10; traduzimos)¹¹, a velocidade está intimamente ligada à vantagem e à eficiência, quanto mais rápido se entendem os dados, mais vantagem adquire-se no mercado competitivo entre as empresas, e torna possível chegar a uma análise eficiente.

As características de *big data* não ficam restritas aos três Vs, anteriormente abordados. Outros autores como Zikopoulos et al. (2012) dizem que o *big data* caracteriza-se por cinco aspectos: *volume*, *velocidade*, *variedade*, *veracidade*, *valor*.

A *veracidade*, segundo Zikopoulos et al., acaba sendo sinônimo de qualidade e confiabilidade dos dados, fazendo com que as análises de *big data* possam oferecer *insights* confiáveis. O *valor* da nova tecnologia está na vantagem competitiva que as análises de *big data* podem conferir às empresas duas vezes mais chances de ultrapassar seus concorrentes na indústria (segundo resultados de um estudo chamado *The New Intelligent Enterprise* realizado pelo *Institute of Business Value* em parceria com o *Sloan Management Review* do *Massachusetts Institute of Technology* (MIT) (ZIKOPOULOS et al, 2012).

Ainda sobre o conceito de *big data*, Thomas Davenport (2014, p. 1), afirma que

⁹ Sobre a definição de dados estruturados e desestruturados, Judith Hurwitz et al. explicam que os primeiros, “[...] se referem a dados que possuem um formato e comprimento definido. Exemplos de dados estruturados incluem números, datas e grupos de palavras e números chamados *strings* (por exemplo, o nome de um cliente, endereço e assim por diante)”, enquanto que os dados desestruturados “são dados que não seguem um formato específico. [...] São, realmente, a maioria dos dados, [...] até recentemente a tecnologia não suportava fazer sua manipulação, exceto armazená-los ou analisa-los manualmente.” (HURWITZ, Judith et al. **Big data for dummies**. Rio de Janeiro, Alta Books, 2015, p. 26, 29)

¹⁰ “To clarify, all data has some structure; when we refer to unstructured data, we are referring to the subcomponents that don’t have structure, such as the freeform text in a comments field or the image in an auto-dated picture.” (ZIKOPOULOS, et al, op. cit., 2012, p. 10; traduzimos).

¹¹ “We define velocity as the rate at which data arrives at the enterprise and is processed or well understood.” (ZIKOPOULOS et al., op. cit 2012, p. 10; traduzimos).

[...] *big data* é um termo genérico para dados que não podem ser contidos nos repositórios usuais; refere-se a dados volumosos demais para caber em um único servidor; não estruturados demais para se adequar a um banco de dados organizado em linhas e colunas; ou fluidos demais para serem armazenados em uma *data warehouse* estático. Embora o termo enfatize seu tamanho, o aspecto mais complicado do *big data*, na verdade, evolui sua falta de estrutura.

Como visto, o *big data* é um grande volume de dados que permite sua coleta, tratamento, armazenamento e reutilização dos dados, inclusive, dados pessoais sendo que a mudança de escala leva a uma mudança de estado. São tantos os dados e de variáveis características que podem não parecer informações pessoais explícitas, mas que, correlacionados, reordenados ou reidentificados¹², após os processos de análises com o *big data*, podem facilmente dizer a quem se referem ou levar ao conhecimento de detalhes íntimos da vida de uma pessoa.

Dois fatos podem constituir a premissa deste tópico. O primeiro foi em 2009, quando o vírus H1N1 assolou o mundo. O *Google* conseguiu, a partir da combinação de centenas de bilhões de termos de busca, identificar o surgimento de um surto de gripe, de forma precisa, em tempo real gerando respostas mais rápidas e eficientes a órgãos oficiais como a Organização Mundial de Saúde (OMS), enquanto que os Centros de Controle e Prevenção de Doenças, em meio à pandemia, estavam sempre defasados em uma ou duas semanas em relação aos dados sobre os novos casos emitidos pelos médicos. O outro foi pela Lei de Etzioni. Através de uma análise da variação de preços das passagens aéreas, pode-se prever o melhor dia e horário para realizar a compra de um bilhete aéreo com o menor custo, dando assim um significativo poder econômico para o consumidor. É nítido que o *big data* é uma ferramenta que, em si, dita novas maneiras de agir, e interfere diretamente na compreensão e reestruturação da sociedade (MAYER-SCHONBERGER; CUKIER, 2013).

O fenômeno sociotécnico em questão está interligado à mudança de escala dos bancos de dados que aumentam a cada dia. Um significativo exemplo é o provedor *Google*, que processa mais de 24 *petabytes* de dados por dia, volume milhares de vezes maior que todo o acervo impresso da Biblioteca do Congresso dos Estados Unidos (CHEDE, 2013).

É inquestionável que a internet provocou mudanças radicais ao acrescentar a comunicação aos computadores. Nessa mesma perspectiva, o *big data* mudará aspectos fundamentais da vida ao lhe dar uma dimensão quantitativa que ela jamais teve (MAYER-SCHONBERGER; CUKIER, 2013). Uma razão evidente que confirma tal assertiva é o valor

¹² O processo de reidentificação refere-se ao processo de análise e tratamento de dados anônimos que, a partir das análises de *big data*, podem revelar informações pessoais, este assunto é abordado adiante nos itens 4.7 e 4.8.

científico e social da nova tecnologia, bem como sua representação como nova fonte de valor econômico e inovação.

Considerando que a finalidade do *big data* está em analisar volume de dados para convertê-los em conhecimento, inovação e valor (DAVENPORT, 2014). A partir de sua abordagem que busca interpretar grande quantidade de dados para extrair novas ideias, é inegável que traga consigo um avanço para diversos setores, dos negócios, às ciências e saúde, governo, educação, e todos os demais setores da sociedade que possam dele se apropriar.

Contudo, o valor da informação não se encontra somente em um propósito (hipoteticamente) primário, com vistas a resultar somente em benefícios para os indivíduos. A partir das análises de *big data*, as informações depositadas nos bancos de dados passam a ter usos secundários. Essas análises referem-se ao processo de correlações dos dados que acaba por revelar informações pessoais, seja o comportamento do indivíduo como consumidor, suas condições de saúde e até atividades ilegais. Em linhas gerais, o *big data* buscar desvendar para as empresas, o comportamento, preferências e aversões dos indivíduos, estes na condição de cliente-consumidor, a fim de possibilitar uma maior segmentação de anúncios, produtos e serviços com base nesse conhecimento (DAVENPORT, 2014).

Os riscos, portanto, passam a residir na capacidade de captação de dados pessoais¹³, assim como na utilização dos dados coletados e armazenados pelo *big data*, que, explicitamente, podem não parecer dados pessoais, mas implicitamente são quando submetidos às análises do *big data* que correlacionam os dados coletados fazendo com que estes transmitam informações pessoais sobre os indivíduos. Assim, une-se a informação coletada, e confere-se valor à informação para propósitos secundários. Diante todo o processo de captação dos dados e correlação, os indivíduos não podem consentir ou optar por não conceder suas informações, pois sequer sabem com exatidão quais os seus dados estão sendo coletados e para quais estão sendo destinados os usos secundários.

As oportunidades de *big data* centram-se, em sua maioria, em produtos, serviços e clientes. Davenport lista inúmeras empresas e exemplos que se evidenciam essas oportunidades:

¹³ “A *Google* coleta praticamente todos os dados que puder sobre os clientes e seu comportamento e às vezes, coleta dados que não deveria. Um bom exemplo disso foi o projeto de mapeamento do *Google StreetView*, que também coletou dados de redes *wi-fi* não criptografadas ao passar pelas cidades. Essas empresas têm recursos suficientes para se darem ao luxo de coletar dados sem qualquer propósito em vista - embora elas pudessem ter ainda mais sucesso se passassem na finalidade antes da coleta dos dados.” (DAVENPORT, op. cit., 2014, p. 165)

Por exemplo, Reid Hoffman, cofundador e presidente do conselho do LinkedIn, montou, com seus cientistas de dados, uma equipe de linha de produto que desenvolveu funcionalidades como “Pessoas que talvez você conheça”, “Grupos que podem ser de seu interesse”, “Vagas que podem ser de seu interesse”, “Quem viu seu perfil”, entre várias outras. A GE se concentra no big data principalmente para melhorar seus serviços e já está usando a ciência de dados para otimizar os contratos de serviço e os intervalos de manutenção para seus produtos industriais. A Google, é claro – a empresa suprema do big data -, usa os cientistas de dados para refinar seus algoritmos centrais de busca e anúncios. A Zynga usa cientistas de dados para orientar o desenvolvimento de games e produtos relacionados com os clientes. A Netflix criou o famoso Netflix Prize para a equipe de cientistas de dados que conseguisse otimizar mais as recomendações de filmes para os clientes (DAVENPORT, 2014, p. 15).

O que as empresas e grandes corporações, que fazem uso das análises de *big data*, têm em comum? Além de ter facilmente à sua disposição os meios de criação de bens e serviços, têm o “conhecimento sobre o conhecimento”, ou seja, a partir das análises de *big data* elas sabem como encontrar, coletar e correlacionar, para, então, direcionar seus bens e serviços aos seus clientes. Todas essas vantagens na era da informação são tão importantes quanto saber quais são as informações.

Na conjuntura da sociedade da informação em que as liberdades individuais são colocadas em risco em razão do surgimento das novas tecnologias, o direito não pode se omitir a esse cenário, como ensina o eminente professor Antonio Enrique Pérez Luño (2006, p. 89),

En un mundo interdependiente, en el seno de sociedades interconectadas, la garantía de los derechos cívicos, se halla en directa conexión, para bien o para mal, con los procesos que definen su instalación tecnológica. El estudio actual de los derechos humanos no puede omitir esa referencia contextual, ni puede abdicar del juicio crítico de sus implicaciones. Se trata de lograr que los desarrollos tecnológicos no menoscaben ni se alcancen a costa de las libertades cívicas. Por ello, las reflexiones interdisciplinarias tendentes a establecer un diálogo fluido entre el universo tecnológico y la esfera de los derechos de los ciudadanos se han hecho cada vez más perentorias.

Por essas razões, diante do valor científico, social e econômico do *big data*, assim como os riscos que recaem sobre os dados no processo de coleta, tratamento e usos desses dados, emerge a necessidade de se discutir a efetividade da proteção jurídica dos dados pessoais como forma de resguardar a privacidade, indispensável no contexto das novas tecnologias.

4 CONSIDERAÇÕES FINAIS

Diante do que foi exposto no presente trabalho, evidenciou-se que as inovações tecnológicas trazem consigo valores econômicos, sociais e científicos para a sociedade, materializadas principalmente a partir da produção de bens e serviços. O *big data*, como visto, surgiu nesse contexto como ferramenta com aptidão para extrair e revelar novas ideias a partir da grande quantidade de dados que são disponibilizadas em escalas sem precedentes na internet, à disposição das grandes corporações. Tem-se que não somente é crescente o uso da internet pelas pessoas, como estas são cada vez mais vulneráveis, em decorrência da ausência de conhecimento sobre quais informações diariamente são disponibilizadas *online*, direta ou indiretamente, ao utilizar os serviços *online*, ao mover-se na sociedade da informação. Emerge, portanto, a necessidade de uma tutela efetiva dos dados pessoais, apta a produzir no mundo dos fatos os efeitos pretendidos pela legislação, aproximando o normativo e a realidade social.

Com a transição da sociedade disciplinar para sociedade de controle, respectivamente objeto dos estudos de Michel Foucault e Gilles Deleuze, foi possível demonstrar que ambas compartilham o atributo da ‘vigilância’ do indivíduo, diferenciando-se apenas nos mecanismos como cada um procede, estando para a sociedade de controle - a que vivemos - os bancos de dados automatizados como o projeto arquitetônico de vigilância pós-moderno, uma reinvenção do panóptico utilizado por Michel Foucault para descrever a sociedade disciplinar.

Os bancos de dados automatizados, devido ao avanço tecnológico, tornaram possível a transição da informação dispersa para organizada; atrelado a isso, através do *big data*, que tem como características o volume, a variedade e velocidade, no processo de armazenamento, coleta e tratamento, respectivamente, o indivíduo tem sido colocado em uma situação de vulnerabilidade representada pela ausência de controle sobre suas informações pessoais.

REFERÊNCIAS

BAUMAN, Zygmunt. **Globalização: as consequências humanas**. Rio de Janeiro: Jorge Zahar Ed., 1999.

BAUMAN, Zygmunt. **Vigilância líquida: diálogos com David Lyon**. Tradução Carlos Alberto Medeiros. Rio de Janeiro: Jorge Zahar Editor, 2014, p. 88.

BRAGA, S.; VLACH, V. Os usos políticos da tecnologia, o biopoder e a sociedade de controle: considerações preliminares. **Scripta Nova. Revista electrónica de geografía y ciencias sociales**. Barcelona: Universidad de Barcelona, 1 de agosto de 2004, vol. VIII, núm. 170(42). Disponível em: <<http://www.ub.es/geocrit/sn/sn-170-42.htm>>. Acesso em 14 ago. 2018.

- CASTELLS, Manuel. **A sociedade em rede**. 9. ed. São Paulo: Paz e Terra, 2006.
- CHEDE, Cezar. **Big data**: expectativas, benefícios e barreiras, 2013. Disponível em: <https://www.ibm.com/developerworks/community/blogs/ctaurion/entry/big_data_expectativas_beneficios_e_barreiras?lang=en> Acesso em 24 ago. 2018.
- CHEN, C.L, Philip; ZANG, Chun-Yang. Data-intensive applications, challenges, techniques and technologies: A survey on Big Data, *Inform. Sci.*, v. 275, 10 ago. 2014, p. 3, traduzimos. Disponível em: <<http://dx.doi.org/10.1016/j.ins.2014.01.015>>. Acesso em 21 ago. 2018.
- CHEVITARESE, Leandro; PEDRO, Rosa Maria Leite Ribeiro. Da sociedade disciplinar à sociedade de controle: a questão da liberdade por uma alegoria de Franz Kafka, em O Processo, e de Phillip Dick, em Minority Report. **Estudos de Sociologia. Rev. do Prog de Pós-graduação em Sociologia da UFPE**, v. 8, n. 12. 2002, pp. 129-162.
- COSTA, Rogério da. Sociedade de controle. **São Paulo Perspec.** São Paulo, v. 18, n. 1, p. 161-167, mar. 2004. Disponível em <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-88392004000100019&lng=pt&nrm=iso>. Acesso em 15 ago. 2018.
- CRESWELL, J. W. **Projeto de pesquisa**: método qualitativo, quantitativo e misto. 3. ed. Porto Alegre: Artmed, 2010.
- DAVENPORT, Thomas H.. **Big data no trabalho**: derrubando mitos e descobrindo oportunidades. Tradução Cristina Yamagami. Rio de Janeiro: Elsevier, 2014.
- DELEUZE, G. *Post-scriptum* sobre as sociedades de controle. In: _____. **Conversações (1972-1990)**. São Paulo: Ed. 34, 1992.
- ECO, Umberto. **Apocalípticos e integrados**. São Paulo: Perspectiva, 2001.
- FEIJÓ, Bruno Vieira. A revolução dos dados. **Revista Exame PME – Pequenas e Médias Empresas**, São Paulo, p. 30-43, set. 2013.
- FOUCAULT, M. **Vigiar e punir**: nascimento da prisão. 38. ed. Petrópolis (RJ): Vozes, 2010.
- HURWITZ, Judith et al. **Big data for dummies**. Rio de Janeiro, Alta Books, 2015.
- MAYER-SCHONBERGER, Viktor; CUKIER, Kenneth. **Big data**: como extrair volume, variedade, velocidade e valor da avalanche de informação cotidiana. Tradução Paulo Polzonoff Junior. Rio de Janeiro: Elsevier, 2013.
- MOREY, Timothy; FORBATH, Theodore; SCHOOP, Alisson, Dados dos consumidores: modelos de transparência e confiança. **Harvard Business Review Brasil**, São Paulo, fevereiro 2016.
- NEVES, Cláudia E. Abbês Baêta. Sociedade de Controle, o neoliberalismo e os efeitos de subjetivação. In: SILVA, André do et al. (Org.). **Subjetividade**: questões contemporâneas. São Paulo: Hucitec, 1997.

PÉREZ LUÑO, A-E. **La terceira generación de derechos humanos**. Cizur Menor (Navarra): Thomson-Aranzadi, 2006.

RUSSOM, Philip. **Big data analytics**. Renton: TDWI Research, 2011.

SIBILIA, Paula. **O homem pós-orgânico: o corpo, subjetividade e tecnologias digitais**. Rio de Janeiro: Relume Dumará, 2002.

TÓTORA, Silvana. Democracia e sociedade de controle. **Verve**, 10: 237-261, 2006, p. 239. Disponível em: <<http://revistas.pucsp.br/index.php/verve/article/viewFile/5441/3888>>. Acesso em 14 ago. 2018.

ZIKOPOULOS, P; DE ROOS, D; PARASURAMAN, K; DEUTSCH, T; GILES, J; CORRIGAN, D. **Harness the power of Big Data-** The IBM Big Data Platform. Emeryville: McGraw-Hill Osborne Media, 2012.