Revista de Direito, Governança e Novas Tecnologias

Organização Comitê Científico
Double Blind Review pelo SEER/OJS
Recebido em: 12.12.2017
Aprovado em: 30.12.2017

MINERAÇÃO DE DADOS E ANÁLISE PREDITIVA: REFLEXÕES SOBRE POSSÍVEIS VIOLAÇÕES AO DIREITO DE PRIVACIDADE NA SOCIEDADE DA INFORMAÇÃO E CRITÉRIOS PARA SUA ADEQUADA IMPLEMENTAÇÃO À LUZ DO ORDENAMENTO BRASILEIRO

Cristiano Colombo[1]
Eugênio Facchini Neto[2]

Resumo: O estudo analisa o tema da mineração de dados, da coleta à análise preditiva, refletindo sobre possíveis violações ao direito de privacidade, à luz dos princípios vigentes na ordem jurídica nacional, mas atento também à experiência do direito comparado. A mineração tem como matéria-prima os dados fornecidos pelos internautas (input), gerando uma fonte geradora de conhecimento (output), impactando na tomada de decisões. Procurou-se analisar a temática avaliando os dois principais modelos de proteção de dados (norte-americano e europeu). Quanto à metodologia, a abordagem da pesquisa foi teórica, exploratória e descritiva. Como procedimento técnico, utilizou-se a pesquisa bibliográfica nacional e estrangeira.

Palavras-chave: Sociedade da Informação – Mineração de Dados – Proteção de Dados – Direito de Privacidade – Big Data.

DATA MINING AND PREDICTIVE ANALYTICS: REFLECTIONS ON POSSIBLE VIOLATIONS OF PRIVACY LAW IN THE INFORMATION SOCIETY AND CRITERIA FOR THEIR PROPER IMPLEMENTATION IN THE LIGHT OF THE BRAZILIAN PLANNING

Abstract: The study analyzes the data mining subject, from collection to predictive analysis, reflecting on violations of the right to privacy, according to the current principles in the national legal order and also the experience of comparative law. The raw material of data mining was provided by the Internet users (input), creating a knowledge-generating source (output), impacting the decision-making process. The theme was analyzed by evaluating the two main models of data protection (North American and European). As for the methodology, the research approach was theoretical, exploratory and descriptive. As a technical procedure, it was used national and foreign bibliographic research.

[1] Pós-Doutorando em Direito pela Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS). Doutor em Direito pela Universidade Federal do Rio Grande do Sul (UFRGS). Mestre em Direito pela Universidade Federal do Rio Grande do Sul (UFRGS). É professor da Faculdade de Direito da Instituição Educacional São Judas Tadeu, Centro Superior Cachoeirinha (CESUCA). Endereço eletrônico: cristiano@colomboadvocacia.com.br.

[2] Doutor em Direito Comparado (Florença/Itália), Mestre em Direito Civil (Universidade de São Paulo). Professor dos Cursos de Graduação, Mestrado e Doutorado em Direito da PUC/RS. Professor e ex-diretor da Escola Superior da Magistratura/AJURIS. Desembargador no Tribunal de Justiça do Rio Grande do Sul/Brasil. Endereço eletrônico: facchini@tj.rs.gov.br.

Keywords: Information Society - Data Mining - Data Protection - Privacy Law - Big Data.

1 INTRODUÇÃO

O presente estudo tem por objeto o tema da chamada mineração de dados, da coleta à análise preditiva, buscando refletir acerca de possíveis violações ao direito à privacidade, bem como investigar critérios para sua adequada implementação, à luz do direito brasileiro. A expressão "mineração de dados" representa para o contexto das organizações, a partir de dados fornecidos pelos próprios internautas (input), no mundo digital, uma fonte geradora de conhecimento (output), impactando diretamente na tomada de decisões pelos seus gestores. Opera-se a análise preditiva, quando se busca prever, antecipar situações, projetar eventos futuros, através de algoritmos computacionais, sem a interação humana. A temática é de grande relevo, na medida que, cada vez mais, aumenta o volume de dados pessoais que estão na rede mundial de computadores, havendo necessidade de definições e regulações acerca da matéria.

O primeiro capítulo apresenta o tema da mineração de dados. Sua primeira parte aborda o crescimento do volume de dados pessoais no mundo virtual, bem como o desenvolvimento do fenômeno da coleta de dados. Na segunda parte, aprofundar-se-á o estudo sobre o significado e alcance da análise preditiva. O segundo capítulo abordará a delicada questão de possíveis violações ao direito à privacidade no âmbito da coleta de dados virtuais, bem como será analisada como se dá a proteção de tais dados em nível mundial. A última parte deste capítulo envolverá os possíveis critérios para uma regular e legítima implementação da mineração de dados no ordenamento jurídico nacional.

Quanto à metodologia, a abordagem da pesquisa será teórica, exploratória e descritiva. Os procedimentos técnicos envolveram pesquisa bibliográfica nacional e estrangeira sobre os temas tratados.

2 MINERAÇÃO DE DADOS: DA COLETA À ANALISE PREDITIVA

2.1 DA COLETA DE DADOS

Com a Web 1.0, os usuários da rede mundial de computadores já experimentavam um grande fluxo de circulação de dados, em tempo real, vencendo os limites geográficos. Contudo, nesta primeira fase, comportavam-se passivamente, como leitores de jornais impressos ou

observadores de *outdoors* publicitários em vias públicas. A mensagem emitida era unidirecional, cujo conteúdo era elaborado por empresas especializadas, que tinham como foco o gosto da maioria (PAGALLO, 2009, p. 705).[3] Não era comum a existência de ferramentas específicas de aferição imediata de opinião como "curtir", "não gostar" ou, "amar", acerca de determinado conteúdo.

O crescimento progressivo de dados transferidos pelos próprios internautas para a rede mundial de computadores operou-se com o surgimento da Web 2.0, passando o usuário de mero consumidor para "prossumidor", ou seja, produtor de conteúdo no ciberespaço (AZAMBUJA, 2012, p. 673). E, nesse ponto, novo viés comportamental se revela fortemente nas redes sociais, em que muitas pessoas passaram a desenvolver um perfil ativo, opinativo, concordando e discordando acerca das mais diversas temáticas, revelando seus posicionamentos e preferências. A internet registra, de forma indelével, os dados, como por exemplo: uma corrida pelo aplicativo Uber ou Cabify, possibilitando recuperar os minutos em que transcorreu, a descrição exata do trajeto, o quanto foi pago, de que modo - dinheiro ou cartão - e a avaliação do serviço e do usuário; se uma pessoa viajou ou não, quais os países que conheceu e quanto tempo foi a sua estada em cada um deles; os seus hábitos de consumo, se comeu sushi ou prefere pizza; de que serviços faz uso; que músicas e filmes são de seu agrado (SCHWAB, 2016, p. 12). Da mesma forma, num ambiente competitivo como o que vivemos, todos os agentes econômicos passaram a agir da mesma forma, aumentando exponencialmente o volume de dados que passaram a ser disponibilizados na rede. Esse processo de modificação comportamental desencadeado pela tecnologia foi descrito com precisão por Marquesone (2016, p. 107-116):

> Suponha que estamos em 1996. Ao acordar, desligo meu despertador e me preparo para ir ao trabalho. Ao sair de casa, meu telefone fixo toca e, ao atender, a secretária da empresa em que trabalho me avisa que estou atrasada para a reunião que havia começado a uma hora. Corro para pegar minha agenda dentro da bolsa e vejo que de fato havia marcado a reunião para aquele horário. Peço desculpas à secretária e aviso que irei rapidamente para a empresa. Arrumo-me às pressas e saio de casa na expectativa que um táxi apareça rapidamente, para que eu possa chegar o quanto antes na reunião. Por sorte, um taxista aparece em 10 minutos. Chego na empresa, porém percebo que esqueci de levar os relatórios que havia elaborado para apresentar aos gerentes. E agora? Ligo para meu marido que está em casa e peço para ele me enviar uma cópia via fax. Assim ele faz, e consigo finalmente participar da reunião. Atualmente, é comum usarmos nosso smartphone desde o primeiro instante em que acordamos, por meio de um alarme com nossa música favorita e por intervalos de tempos pré-determinados. Nosso smartphone também pode nos avisar antecipadamente o horário de uma reunião, para que assim possamos evitar esquecimentos. Enquanto tomamos café, podemos solicitar um serviço de transporte

[3] O autor opõe a passividade inicial versus a interatividade experimentada posteriormente.

> de passageiros por meio de um aplicativo. Se necessitamos de um documento que não esteja conosco, podemos facilmente acessar a internet e busca-lo em um serviço de computação em nuvem para armazenamento de dados.

Na descrição acima, percebe-se que as pessoas passaram a depender do mundo virtual, em face da maior *disponibilidade* dos dados e na possibilidade de acesso remoto aos mesmos, independentemente de onde quer que se esteja, sem a necessidade da presença física do agente. Outro fator é a *integridade* dos mesmos, eis que ao alcance do usuário, em seu inteiro teor, oferecendo a mesma qualidade como se estivessem fisicamente sendo avaliados. E, por último, a *confidencialidade*, no sentido de que os provedores buscam preservar o seu conteúdo do acesso de terceiros (BARROS, 2015, p. 687).

Como refere José de Oliveira ASCENSÃO (2001: p. 84), essa abertura mundial das comunicações, que permite acesso e circulação livre de um volume ilimitado de informações, "pode ser comparada à abertura mundial dos portos, em que o Reino Unido se empenhou no século passado. A diferença está em que para esse efeito não foi até agora utilizada a canhoneira". Neste sentido, cumpre esclarecer que a expressão "mineração de dados" revela per se que a coleta, o tratamento e a utilização de dados no mundo virtual, representa para o contexto das organizações, assim como se dá na extração de minérios, uma fonte geradora lucratividade, impactando diretamente na tomada de decisões pelos seus gestores. Para a perfeita compreensão deste processo, que se inicia pela extração de dados em seu estado bruto do meio ambiente digital, há que se analisar três conceitos ligados ao fenômeno: dado, informação e conhecimento. Segundo Leandro Augusto da Silva, tem-se que:

> O dado é um fato, um valor documentado ou um resultado de medição. Quando um sentido semântico ou um significado é atribuído aos dados, gera-se informação. Quando estes significados se tornam familiares, ou seja, quando um agente os aprende, este se torna consciente e capaz de tomar decisões a partir deles, e surge o conhecimento (SILVA; PERES; BOSCARIOLI, 2016, p. 384-386).

Exemplificativamente, o número de curtidas pelos internautas sobre um vídeo musical postado na rede mundial de computadores é meramente um dado. Atribuir a este número a compreensão que revela a plena aceitação do músico ou, quem sabe, a sua rejeição, é uma informação. Já promover ou não a contratação do artista para ligar o nome e a imagem à determinada marca ou organização é conhecimento, pois representa tomada de decisões.

Os dados coletados na web podem ser classificados em estruturados e não estruturados. Os dados estruturados são aqueles que estão organizados, ou seja, estão tabulados em linhas e colunas. É o que se verifica, por exemplo, na coleta de dados decorrentes do Portal da Transparência, que identifica o servidor público federal, seu cargo e sua remuneração. Neste

caso, não há necessidade maior tratamento destes dados para que se possa a atingir o estágio da informação e, por último, do conhecimento. Por sua vez, dados não estruturados são os textos, imagens, vídeos e sons. Dessa forma, é preciso maior custo, tempo e expertise para interpretação destes dados. É o que ocorre com os comentários escritos por internautas, em geral, em blogs e em redes sociais que deverão ser tratados, utilizando-se técnicas de identificação de palavras, números de ocorrências, bem como o reconhecimento do idioma que está escrito, permitindo identificar a origem geográfica dos usuários, disso tudo retirando informações. Atualmente, a mineração de dados se vale tanto dados estruturados como não estruturados, o que significa dizer que, mesmo o compartilhamento de símbolos, bandeiras, cores, sinais, distintivos, ainda que não tabulados, são matérias-primas extremamente valiosas, pois fornecem informações úteis para agentes que desenvolveram a capacidade de atribuir a tais dados interessantes e potencialmente rentáveis significados mercadológicos.

2.2 DA ANÁLISE PREDITIVA

Ocorre que a mineração de dados não se resume ao ato de captura dos dados estruturados ou não, como um fim em si mesmo. Trata-se, na verdade, de um processo com etapas a serem observadas, com uma finalidade específica. Nas palavras de SILVA; PERES; BOSCARIOLI (2016, p. 495):

> De forma simplificada, a mineração de dados pode ser definida como um processo automático ou semiautomático de explorar analiticamente grandes bases de dados, com a finalidade de descobrir padrões relevantes que ocorrem nos dados e que sejam importantes para embasar a assimilação de informação importante, suportando a geração de conhecimento.

Em verdade, é um olhar para o passado, a partir dos dados que estão diante do gestor, para fundamentar decisões futuras. Esta forma de agir não é recente. No entanto, os meios tecnológicos aprimoraram a coleta, o tratamento e a utilização dos dados. A atualidade do tema se verifica por estar sendo aplicada amplamente por gigantes do setor de varejo, como se depreende da reportagem publicada no Valor Econômico, em 26 de julho de 2017:

> Um dos desafios para os profissionais das áreas de planejamento é o chamado "efeito chicote", ou a falta de sincronia entre os pedidos das lojas e os centros de distribuição da Kopenhagen e Brasil Cacau, com mais de 900 pontos de venda em todo o país, entre lojas próprias e franquias. 'O atendimento de tantas lojas gera uma complexidade enorme de planejamento, principalmente por se tratar de produtos altamente perecíveis e que apresentam grande sazonalidade de venda', afirma Fernando Vichi, vice-presidente financeiro, logística e TI do grupo. A rede sofria com o desperdício de mercadorias com vencimento do prazo de validade e falta de produtos

> na prateleira, gerando a contratação de fretes extras e a reprogramação da produção na fábrica. Com a adoção do sistema de análise de dados da Tevec foi possível acompanhar os principais indicadores de abastecimento, identificando a quantidade ideal sem interação humana no processo. [...] Na rede de lavanderias 5àsec, informações colhidas dos clientes por um sistema de big data da Disruptiva avalia os hábitos desses consumidores para fazer previsões de vendas e orientar promoções (MAHLMEISTER, 2017).

Como se vê, não há dúvidas que a expressão "previsões" se apresenta como ponto central e de maior interesse no trecho da notícia colacionada, sendo que poderia ser plenamente substituída no texto por "predições", na linha do estudo ora desenvolvido. Na análise preditiva, objeto do estudo, busca-se prever, antecipar situações, projetar eventos futuros, a partir de algoritmos computacionais, ou seja, sem a interação humana. Para a adequada compreensão, parte-se de um exemplo extremamente corriqueiro que pode ser experimentado por qualquer usuário da internet, ao redigir buscas no "pesquisar Google":

> Em essência, o algoritmo foi capaz de completar ou 'prever' a situação mental dos sujeitos e esses processos de pensamento são muito parecidos com os usados em motores de busca na internet ou programas de mensagens de texto em telefones celulares, os quais podem antecipar e completar uma frase ou pedido antes que o usuário termine de digitar" (DAQUINO, 2017).

Em sendo assim, automaticamente, através da aplicação de algoritmos, sem a interação humana, atinge-se a predição, a projeção do evento futuro:

> A análise preditiva é mais complexa do que a descritiva e a diagnóstica. Elas exigem o uso de grandes conjuntos de dados históricos para permitir assim prever a classe de um conjunto de observações, baseando-se na similaridade de observações classificadas no passado. [...] O propósito aqui é que o algoritmo seja capaz de se adaptar de acordo com os parâmetros recebidos por ele, de forma que sua capacidade de predição e otimização seja feita automaticamente (MARQUESONE, 2016, p. 2381).

Dessa forma, através de técnicas realizadas por algoritmos computacionais, que receberão como entrada (*input*) dados que são fatos da vida real, como idade, profissão, sexo, preferências, níveis de sedentarismo, poderão devolver conhecimento (*output*), como, por exemplo, a um plano de saúde, prevendo custos e internações com determinado paciente ou população de usuários (SILVA; PERES; BOSCARIOLI, 2016, p. 532). Mas, o que é um algoritmo?

> Um algoritmo nada mais é do que uma receita que mostra passo a passo os procedimentos necessários para a resolução de uma tarefa. Ele não responde a pergunta "o que fazer?", mas sim "como fazer". Em termos mais técnicos, um algoritmo é uma sequência lógica, finita e definida de instruções que devem ser seguidas para resolver um problema ou executar uma tarefa (PEREIRA, 2017).

Nessa mesma linha, outra aplicação interessante seria a utilização destas técnicas em um restaurante, onde seria possível gerar uma função a partir de uma regra de associação, ligando o consumo de determinados pratos com vinhos que melhor se harmonizam. E, com o aparecimento de um novo prato, a partir de seus ingredientes (*input),* descobrir qual o vinho passaria a ser mais consumido com ele (*output)* (SILVA; PERES; BOSCARIOLI, 2016, p. 532). Apresentados os principais elementos e conceitos acerca da mineração de dados, ora se passa a refletir sobre eventuais violações do direito de privacidade que decorrem do uso de tais ferramentas, bem como as cautelas que devem ser observadas para que se possa corretamente implementá-las, à luz de nosso ordenamento jurídico brasileiro.

3 REFLEXÕES SOBRE POTENCIAIS VIOLAÇÕES AO DIREITO DE PRIVACIDADE NA SOCIEDADE DA INFORMAÇÃO E CRITÉRIOS PARA SUA PRESERVAÇÃO, NO MUNDO VIRTUAL

3.1 REFLEXÕES SOBRE POTENCIAIS VIOLAÇÕES AO DIREITO À PRIVACIDADE NA SOCIEDADE DA INFORMAÇÃO

Em 16.12.1999, a conhecida revista britânica *The Economist* publicou uma reportagem intitulada "Living in the Global Goldfish Bowl" (Vivendo no aquário global). Tratava-se de um experimento, segundo o qual um jornalista do seu staff havia contratado um investigador particular, encarregando-lhe de fazer uma investigação de uma pessoa, fornecendo-lhe apenas o seu nome (que era o do próprio jornalista). Pediu-lhe, também, que conduzisse a investigação sem se afastar do seu escritório. No espaço de uma semana, a partir apenas desse dado, o investigador conseguiu levantar um impressionante número de informações sobre aquela pessoa, dentre as quais o seu salário, seus telefones, o valor de sua casa, nome e endereço de seus parentes próximos, nome do atual e anteriores empregadores, bem como de sua atual parceira e de sua anterior parceira (REHM: 2003, p. 374). E isso há quase vinte anos atrás, quando ainda o fenômeno da acumulação e circulação de dados não havia atingido as gigantescas proporções de atualmente, especialmente a partir das redes sociais. Hoje, idêntico resultado, com maior detalhamento, seria alcançado no giro de poucas horas.

Não é a primeira vez que direitos ligados à personalidade, como o direito de privacidade, são colocados em xeque diante da evolução tecnológica. Isto se demonstra pelo

famoso episódio que substancialmente inaugurou o debate sobre o direito à privacidade, em nível mundial: tratava-se do surgimento das primeiras máquinas fotográficas instantâneas e seu uso pela chamada 'imprensa marrom', em 1890, o que motivou o ensaio de Warren e Brandeis sobre "The Right to Privacy", consagrando o direito de estar só (PAGALLO, 2008, p. 4).

No espaço jurídico europeu, versão mais antiga do direito à privacidade nos é dada na Resolução 428 (1970) da *Consultative (Parliamentary) Assembly of the Council of Europe,* ao referir que (apud KRZEMINSKA-VAMVAKA/O'CALLAGHAN: 2010, p. 115)

> "The right to privacy consists essentially in the right to live one's own life with a minimum of interference. It concerns private, family and home life, physical and moral integrity, honour and reputation, avoidance of being placed in a false light, non-revelation of irrelevant and embarrassing facts, unauthorized publication of private photographs, protection from disclosure of information given or received by the individual confidentially. Those who, by their actions, have encouraged indiscreet revelations about which they complain latter on cannot avail themselves of the right to privacy".

A esfera da privacidade, no seu sentido clássico, ainda permanece importantíssima. Cada um de nós necessita de um espaço íntimo, longe de olhares e da percepção dos outros, para que possa relaxar e sermos nós mesmos, sem o peso que a observação alheia permanente inerentemente acarreta (CREMER: 2011, p. 62). A importância de tal aspecto foi realçado no relevante caso *Caroline Von Mônaco II,* julgado pela Corte Constitucional alemã em 1999:

> This sphere also offers the opportunity to behave in a way which is not for the public to perceive and the observation or presentation of which by outsiders would be embarrassing or disadvantageous to the person concerned. But essentially it is a sphere in which he has the possibility to be free from public observation and thus from self-control thereby forced upon him. If such spaces of retreat no longer existed, the individual could be psychically overburdened as he would incessantly have to be cautious of his impression on others and of whether he is behaving correctly.

Atualmente, o fenômeno das modernas formas de gigantesca coleta de dados pessoais alterou a visão tradicional da privacidade em vários aspectos. Em primeiro lugar, as questões relacionadas à privacidade, que classicamente envolviam um indivíduo isolado (o clássico *right to be let alone*), envolvem simultaneamente milhões de pessoas, considerando a coleta de dados pessoais de consumidores, contribuintes, pacientes, usuários de todos os tipos de serviços, empregados, clientes, pensionistas, assalariados, ou seja, de todos nós. Em segundo lugar, vários dispositivos são capazes de transmitir informações a nosso respeito – celulares, GPS, cartões de crédito, redes sociais, etc – de forma a se poder reconstituir quem nós somos, por onde circulamos, o que consumimos e o que pensamos. Em terceiro lugar, todas essas informações podem ser utilizadas não só para compreender quem nós somos e o que fazemos,

mas principalmente para influenciar nossas condutas, principalmente enquanto consumidores (SIMITIS: 1987, p. 707). Além disso, o fenômeno deixou de ser territorial para ser global, já que o tratamento de dados passou a envolver elementos transnacionais e globais, envolvendo pessoas localizadas em várias partes do mundo, sujeitas a jurisdições diversas e a diferentes normas de proteção de dados pessoais (POULLET; ASINARI; PALAZZI: 2009, p. 12).

Rodotà (2008: p. 17) aponta que "a distinção entre o direito ao respeito da vida privada e familiar e o direito à proteção dos dados pessoais não é bizantina". Aquele reflete um componente individualista, reduzido substancialmente ao poder de impedir a interferência alheia na vida privada de alguém. Trata-se substancialmente de um direito estático, negativo. Já a proteção de dados envolve o controle dos mecanismos de processamento de dados, do qual participa não só as pessoas diretamente interessadas, mas também uma autoridade pública independente, revelando uma proteção dinâmica e positiva, diante do seu aspecto coletivo.

Na sequência, referiremos dois casos, um europeu e outro argentino, que revelam essas duas possíveis aplicações da noção de privacy, a primeira no sentido clássico de vedação de interferência em aspectos da esfera privada, e a segunda no sentido da proteção de dados pessoais. O primeiro é um caso francês (*Spileers v. SARL Omni Pac*) julgado pela Corte de Cassação francesa, em 12.01.1999. O caso versava um representante comercial chamado Spileers, cujo contrato de trabalho previa que ele devia se mudar com sua família para sua nova área de atuação, no prazo de seis meses da sua transferência. Inicialmente designado para atender a região norte e nordeste da França, ele se estabeleceu, com sua família, em Paris, de onde comodamente podia atender referida área. Posteriormente foi transferido para a região de Montpellier, no sul da França. Na mesma semana ele arranjou acomodações para si naquela cidade, mas se recusou a transferir sua família. Invocando a cláusula contratual, seu empregador o demitiu. Na disputa judicial que se seguiu, as cortes inferiores (*Conseil de Prud'hommes* e, na via recursal, a *Cour d'Appel*) validaram sua demissão, em razão da expressa cláusula contratual. A Corte de Cassação, órgão de cúpula da justiça francesa, reformou as decisões anteriores, invocando o art. 8º da Convenção Europeia dos Direitos Humanos, que garante a qualquer pessoa o direito de livremente escolher onde morar com sua família. Prosseguiu a corte, afirmando que em relações desiguais, seja de natureza pública ou privada, a parte mais forte deve levar em consideração os interesses familiares da parte mais vulnerável (HUNTER-HENIN: 2007, p. 106). Já no interessante caso julgado pela Corte Suprema de Justicia de la Nación Argentina em 24.02.2009 (*caso Halabi v. Estado Nacional*), em que se declarou a

inconstitucionalidade da Ley 25.873, de 2003, e sua regulamentação – referidas normas obrigavam as empresas de telecomunicações a armazenar por dez anos os dados de tráfego e de usuários, com a finalidade de poderem ser utilizados com fins probatórios em processos judiciais -, aquela Corte Suprema encarou a privacidade não só como um bem individual (o direito à vida privada e o de excluir terceiros dessa esfera), como também um bem coletivo (direito à proteção de dados, acesso aos dados pessoais e *habeas data*). De fato, o aspecto coletivo do direito à privacidade pode ser afetado não só com normas, mas também com fatos concretos do setor privado, especialmente políticas empresariais. Caso o Google decida mudar sua política de privacidade, milhões de *logs* podem cair no esquecimento ou serem entregues às autoridades, ou usados para fins de marketing. Se o Facebook resolver alterar suas condições de uso, pode decidir que toda a informação fornecida por seus usuários lhe pertence (incluindo fotos, vídeos, imagens e dados pessoais), caso em que pouco se poderia fazer a nível local. Por outro lado, caso uma forte autoridade nacional (o governo norte-americano, por exemplo), resolver, em nome da segurança e defesa de seu país, interferir no tráfego da internet, através da NSA, isso poderia afetar não só as comunicações de seus cidadãos ou de outras pessoas com cidadãos norte-americanos, mas também comunicações envolvendo pessoas que não são súditas daquele país, em razão da arquitetura da rede mundial de computadores. Portanto, o fenômeno não pode mais ser encarado como uma questão puramente de privacidade individual, em seu estilo clássico (PALLAZI: 2009, p. 53/54).

Esse cenário é realmente preocupante, especialmente quando se toma conhecimento dos dados fornecidos pelo professor Daniel Solove, da George Washington University, a partir de realidade de aproximadamente dez anos atrás (o que permite a ilação de que os números por ele referidos tenham aumentado enormemente). Refere o professor que está surgindo uma nova espécie de atividade empresarial, voltada exclusivamente à coleta e tratamento de dados pessoais. A empresa Catalina Marketing Corporation, com sede na Florida, mantém bases de dados reunindo históricos de compras de supermercados de mais de 30 milhões de famílias. A empresa Aristotle International comercializa uma base de dados de 150 milhões de eleitores registrados, reunindo dados como nome, endereço, telefones, filiação partidária e frequência de votos (lembrando-se que nos EUA, o voto é facultativo). Aristotle combina esses dados com mais 25 classes de informações pessoais, dentre as quais raça, renda, emprego e até modelo do seu carro. Dentre outras finalidades, costuma vender uma lista de pessoas ricas que fazem doações eleitorais, denominada de "Peixe grande". No seu marketing, Aristotle alardeia:

"Pegue seu adversário na bilheteria! Com 'Peixe grande' você pode descobrir quem são os contribuintes da campanha de seu adversário, endereçar-lhes uma correspondência explicando porque não deveriam contribuir para ele". Outra empresa disponibiliza uma aplicação chamada "GeoVoter", que combina aproximadamente 5.000 categorias de dados sobre cada eleitor, para calcular a probabilidade de voto desta pessoa. Já a empresa Wiland Services construiu uma base de dados contendo cerca de 1.000 aspectos envolvendo a conduta de mais de 215 milhões de pessoas. Por último, refere-se que há no mínimo cinco empresas compiladoras de dados que possuam informações sobre praticamente todas as famílias dos Estados Unidos (SOLOVE: 2009, p. 87/88).

Os dois lados da moeda – privacidade clássica e proteção de dados – muitas vezes estão ligados, como é o caso do direito ao esquecimento. A capacidade humana de relembrar, que nos acompanha desde a época das cavernas, permitiu que o homem comparasse, aprendesse e evoluísse. Igualmente importante, porém, é a habilidade humana de esquecer, deixando para trás o peso do passado e permitindo viver o presente de forma mais intensa. Por milênios, a relação entre lembrar e esquecer permaneceu clara. Lembrar é difícil e custoso e os humanos tinham que deliberadamente escolher o que lembrar. O normal era o esquecimento. Na era digital, essa equação se inverteu. Com a facilidade de armazenar um volume impressionante de informações, a memória digital tornou o passado um eterno presente. Lembranças passam a ser eternas e o esquecimento tornou-se exceção (MAYER-SCHÖNBERGER: 2009, p. 196). Dentro deste cenário, adquire renovada importância o direito à identidade pessoal, não só como direito a sermos 'nós mesmos' e o correlato direito ao respeito às nossas próprias escolhas de vida, como também o direito de cada um a uma correta representação exterior de sua própria individualidade. A isso se agrega o direito à privacidade, entendida não só em termos de um *ius excludendi alios* de uma esfera íntima, mas também como direito a não sofrer indevidas interferências externas relativamente à manifestação social de sua própria identidade, confirmando a liberdade de cada um de escolher o que revelar e o que manter só para si (CALIFANO: 2016, p. 200) Portanto, diante de tais reais e atuais ameaças, é preciso observar que o fenômeno ligado à técnica de *data mining*, ou mineração de dados, estejam amalgamados à observância dos direitos que protegem a pessoa humana, em especial o direito à privacidade, que configura simultaneamente um direito fundamental e um direito de personalidade.

O problema é que tais dados, na atualidade, vão muito além de combinações em uma linguagem binária, constituindo-se, em verdade, no próprio *corpo eletrônico* do ser humano:

> O '*corpo elettronico'*, é uma reunião de informações que constroem a nossa identidade, integradas ao corpo físico: a dignidade passa a ser o forte meio para reconstruir a integridade da pessoa (Carta dos Direitos Fundamentais da União Europeia, art. 3), para evitar que a pessoa venha a ser considerada uma espécie de mina a céu aberto, onde qualquer um pode alcançar qualquer informação e assim construir perfis individuais, familiares, de grupo, transformando a pessoa em objeto de poderes externos, que possam falsificá-la, construí-la em conformidade às necessidades de uma sociedade da vigilância, da seleção social, do cálculo econômico (RODOTÀ, 2013, p. 33, Tradução livre do autor).

E, como advertia Stefano Rodotà, os dados pessoais não podem ser considerados produtos de uma "mina ao céu aberto", em que *players* do mundo virtual possam se servir livremente, sem quaisquer limitações, alimentando algoritmos computacionais tendo como única finalidade o lucro. E, dependendo da informação que os dados podem revelar a terceiros, sejam eles estruturados ou não, os mesmos podem ser classificados como dados sensíveis, conforme leciona Laura Schertel Mendes, voltando-se a questões religiosas, de saúde, políticas, étnicas, ou afetas à vida sexual. (2014, p. 217). Nesse sentido, não há como deixar de ser observada a privacidade: "Portanto, vive-se em um ambiente repleto de informações. A infosfera, nesta linha, são "dados e informações destinados a serem partilhados, selecionados, modificados e revistos", devendo a privacidade ser uma espécie de "imunidade" às trocas desconhecidas, indesejadas e não intencionais de informações" (COLOMBO, 2017). O quanto estamos vulneráveis a isso é posto em destaque por Denny Cherry (2014, p. 212):

> Um dos mais notórios problemas, que se tornou manifesto para o público em geral, aconteceu com uma grande varejista. A Target começou a enviar cupons com base em hábitos de compras pessoais que a loja rastreava por meio de programa de fidelidade. Uma cliente específica, que morava com pai, começou a receber cupons de vitaminas pré-natais e suplementos para o bebê. O pai ficou indignado, porque a varejista estava enviando os tais cupons à filha menor de idade; então, foi até a loja mais próxima de sua casa e reclamou com o gerente. A filha, com isso, teve de explicar ao pai que realmente estava grávida. A loja conseguira descobrir a gravidez da adolescente pela análise das compras da cliente na loja, por meio do cartão de programa de fidelidade, o que chamamos de análise de dados.

Ora, no caso acima relatado, a aquisição de produtos na rede da loja acima mencionada configurou-se em uma entrada (*input*), que, em face de nítidas características de uma análise preditiva, por associação, desencadeou o oferecimento de produtos vitamínicos à gestante, bem como a produtos ao bebê (conhecimento). Dessa forma, verifica-se que é possível verificar a ocorrência de violações ao direito de privacidade pela prática da mineração de dados. E, acentue-se que, no episódio colacionado, a violação à privacidade não se deu no dado em si, que foi a aquisição de produtos. O que ofendeu o direito de personalidade foi o resultado obtido pela análise preditiva, enfim, o conhecimento que foi alcançado, através do algoritmo

computacional, que projetou o estado gravídico, a partir dos dados coletados, gerando ofertas direcionadas pela internet. Dessa forma, não somente o dado é sensível, mas, no caso concreto, está-se diante do próprio *conhecimento sensível.* No entanto, tais preocupações, por mais que devam ser levadas a sério, não estão a sugerir a solução simplista de se considerar ilegal a prática da mineração de dados, sugerindo sua vedação. Além de ser impraticável, resultaria em uma grave ofensa ao princípio da liberdade econômica. Ora, de longa data, prestadores de bens e de serviços buscam colher dados de seus clientes para melhor atendê-los, para entender suas necessidades, e, também expandir seus resultados: a mineração de dados é, sem dúvida alguma, uma ferramenta que se enquadra nessa situação. É o que adverte Sérgio Amadeu da Silveira, Rodolfo Avelino, Joyce Souza (2016, p. 217-230):

> Sob a ótica estritamente econômica, o uso de dados pessoais poderia reduzir as assimetrias da informação e contribuir para aumentar a eficiência das transações econômicas nas redes digitais. As corporações podem analisar os dados recebidos dos seus consumidores e organizar estratégias personalizadas para seus produtos e serviços. A tese liberal presente nos discursos das consultorias e dos dirigentes das corporações da economia informacional advoga que a consolidação do mercado de dados pessoais beneficiaria a todos: as empresas colocariam produtos mais adequados e mais compatíveis com a demanda e, ao mesmo tempo, os consumidores seriam alertados sobre as oportunidades de atendimento de acordo com seus interesses. Algoritmos disponíveis nas plataformas online, ao identificarem um certo consumidor, poderiam melhorar a experiência de navegação e de consumo desse indivíduo, uma vez que aquilo que ele mais gosta seria diretamente oferecido, sem perda de tempo, portanto, sem desperdício de recursos econômicos.

Para comprovar o acima dito, pense-se na frequente situação em que, após ter encomendado vários livros sobre determinado tema, no site Amazon, o consumidor começa a receber avisos de que há outros livros disponíveis no mercado envolvendo o mesmo assunto. Por vezes tais informações são preciosas, alertando o potencial consumidor sobre algo que ele decididamente tem interesse e até então ignorava. Portanto, a vedação da coleta de todo e qualquer dado ou predição, além de ser um comando que cairá no vazio da ineficácia legislativa, pois é difícil "enquadrar-se" normativamente o mercado, é desarrazoada. O potencial conflito reclama, isso sim, uma necessária harmonização das práticas e interesses potencialmente contrapostos, com observância dos direitos fundamentais. Nas palavras de Danilo Doneda (2006:12),

> Ocorre que esta atividade requer instrumentos que a harmonize com os parâmetros de proteção da pessoa humana ditados pelos direitos fundamentais, instrumentos que possibilitem aos interessados um efetivo controle em relação aos seus dados pessoais, garantindo o acesso, a veracidade, a segurança, o conhecimento da finalidade para a qual são utilizados (entre outros).

De tal arte, importa destacar que não é o momento de levantar bandeiras contra ou a favor da mineração de dados, que é uma realidade que não se pode refutar. Os esforços devem

ser direcionados à construção de critérios para a sua implementação, harmonizando-a com o ordenamento jurídico pátrio e seus valores mais elevados.

3.2 CRITÉRIOS PARA ADEQUADA IMPLEMENTAÇÃO DA *DATA MINING* À LUZ DO ORDENAMENTO BRASILEIRO

3.2.1 Do Panorama Mundial

O problema do controle dos bancos de dados, em razão de seu potencial uso danoso, vem sendo objeto de preocupação e regulamentação normativa já há várias décadas. Na Europa, a Suécia, em 1973 (Lei 289, denominada *Datalagen*), e a Alemanha, em 1977, foram os primeiros países a se dotarem de leis regulamentadoras dos bancos de dados. Em nível constitucional, coube a Portugal a primazia de incluir no texto de sua constituição o direito de cada cidadão tomar conhecimento dos próprios dados pessoais quando coletado em banco de dados. Diante do caráter lacunoso da sua *Grundgesetz,* coube à Corte Constitucional alemã, em 1983, criar o importante conceito de *direito fundamental à autodeterminação informativa* (*Grundrecht auf informationelle Selbstbestimmung*) ou seja, o poder de acesso e controle dos próprios dados pessoais e o direito de selecionar o que cada indivíduo quer expor de si mesmo aos outros (PAESANI: 2012, p. 35.) A temática da mineração está diretamente vinculada à proteção de dados, que, diante da legislação aplicada em cada país e a existência ou não de uma autoridade oficial competente para a proteção de dados (no âmbito da União Europeia, o novo Regulamento de 2016, que vigorará a partir de 2018, obriga os Estados Membros a se dotarem de tais Autoridades independentes), reserva menor ou maior relevo ao direito à privacidade do cidadão. Em nível mundial, é possível identificar dois grandes modelos de proteção de dados, o norte-americano e o europeu[4], a saber:

> Centram-se ora no indivíduo e nos seus direitos e liberdades (modelo presente no *Privacy Act* dos Estados Unidos da América), ora no controlo da tecnologia e das suas aplicações, ora no controlo institucional (modelo seguido pela União Europeia, todavia actualmente em mudança). Em contraste com a óptica inicialmente privilegiada, centrada na proteção dos direitos individuais, o direito da proteção de dados tem tendido a ser pensado, em particular na Europa, como um instrumento de políticas públicas visando fomentar os serviços e indústrias da informação (GONÇALVES, 2017).

O modelo norte-americano de proteção de dados tem uma estrutura bipartida, eis que

[4] O modelo europeu é seguido, também, pela legislação de vários países latino-americanos, como é o caso da Argentina, Chile, Colômbia, Uruguay, Peru e Paraguay (POULLET; ASINARI; PALAZZI: 2009, p. 11/12).

confere um tratamento ao setor público, com regulamentação mais severa, ao passo que relega o setor privado à autorregulamentação, como leciona Laura Schertel Mendes (MENDES: 2014, p.54). Esta liberdade do setor privado se manifesta na própria política de privacidade do Google, que, de modo geral, informa ao seu usuário como serão coletados e utilizados os dados, sem, no entanto, dar maior especificidade, como se vê de certos trechos de seu documento básico a respeito: a) Utilização dos dados para oferecimento de produtos para a obtenção de resultados "mais relevantes": "Também usamos essas informações para oferecer ao usuário um conteúdo específico, por exemplo, fornecer resultados mais relevantes de pesquisa e anúncios (POLÍTICA..., 2017b); b) Utilização dos dados para melhorar a "experiência do usuário" e "qualidade geral": "Usamos as informações coletadas de cookies e de outras tecnologias, como etiquetas de pixel, para melhorar a experiência do usuário e a qualidade geral dos nossos serviços." (POLÍTICA..., 2017b); c) Utilização dos dados com análise preditiva (propaganda personalizada): "Nossos sistemas automatizados analisam o conteúdo do usuário (incluindo e-mails) para fornecer recursos de produtos relevantes ao usuário, como, por exemplo, resultados de pesquisa e propaganda personalizados e detecção de spam e malware." (POLÍTICA..., 2017b). Portanto, do que se depreende, verifica-se que a Google coleta dados e utiliza a análise preditiva, visto que declara expressamente que "nossos sistemas automatizados analisam o conteúdo do usuário (incluindo e-mails) para fornecer recursos de produtos relevantes ao usuário". Basta saber se todos os usuários estão cientes de tais práticas, e, podem vir a projetar tudo que está sendo feito com seus dados e predições que dela resultam. Ainda, a Google manifesta, de forma expressa, que está lendo todas as mensagens eletrônicas, e, assim, está tendo acesso a dados pessoais sensíveis, gerando, portanto, informações sensíveis e, por conseguinte, está apta a gerar "predições sensíveis", que resultarão na violação de direito de privacidade.

Por sua vez, em relação à União Europeia, foi publicado, em 2016, o Regulamento Geral sobre Proteção de Dados, que passará a ter vigência em maio de 2018. Referido Regulamento recolheu parte da 'herança' da Diretiva de 1995 e avançou em algumas outras direções. A Diretiva 95/46/CE havia nascido da necessidade de garantir aos cidadãos europeus a chamada 'autodeterminação informativa', isto é, o pleno controle sobre seus dados pessoais, embora concedesse alguma flexibilidade aos países integrantes do bloco na transposição normativa para as legislações nacionais. Agora, com o instrumento do Regulamento, todos os países ficaram diretamente vinculados às normas editadas, delas não podendo se afastar. Foram

introduzidos inúmeros institutos que se inscrevem numa perspectiva de caráter responsabilizante. Positivou-se o direito ao esquecimento, na esteira da importante decisão da Corte de Justiça da União Europeia no caso *Google Spain SL e Google Inc. c. Agencia Española de Protección de datos (AEPD), Mario Costeja Gonzáles*, de 13.05.2014. Todos os países membros da U.E. foram compelidos a instituírem/reforçarem suas Autoridades nacionais (espécie de agência administrativa autônoma) com funções/poderes de vigilância, controle e sanção. Importantíssimo foi também a tomada de posição sobre o âmbito de aplicação material do Regulamento, superando o critério do estabelecimento adotado pela Diretiva de 1995. Estão sujeitos ao novo Regulamento e, portanto, subordinados às Autoridades europeias, todos os tratamentos envolvendo dados pessoais de cidadãos europeus, pouco importando se os responsáveis por esses tratamentos são pessoas físicas ou jurídicas não europeus e com sede fora da União Europeia. As novas regras europeias, portanto, atingirão também os colossais impérios da *Information and Communications Technologies* (ICT) como Google, Facebook e Amazon (CALIFANO: 2016, p. 69/71 e 77).

Em seu artigo 6º, assim dispõe sobre a questão da coleta e tratamento de dados:

> 1. O tratamento só é lícito se e na medida em que se verifique pelo menos uma das seguintes situações: a) O titular dos dados tiver dado o seu consentimento para o tratamento dos seus dados pessoais para uma ou mais finalidades específicas; b) O tratamento for necessário para a execução de um contrato no qual o titular dos dados é parte, ou para diligências pré-contratuais a pedido do titular dos dados; c) O tratamento for necessário para o cumprimento de uma obrigação jurídica a que o responsável pelo tratamento esteja sujeito; d) O tratamento for necessário para a defesa de interesses vitais do titular dos dados ou de outra pessoa singular; e) O tratamento for necessário ao exercício de funções de interesse público ou ao exercício da autoridade pública de que está investido o responsável pelo tratamento; f) O tratamento for necessário para efeito dos interesses legítimos prosseguidos pelo responsável pelo tratamento ou por terceiros, exceto se prevalecerem os interesses ou direitos e liberdades fundamentais do titular que exijam a proteção dos dados pessoais, em especial se o titular for uma criança.

Os comentaristas da nova normativa acentuam que o novel Regulamento amplia significativamente o número de operações que passaram a ser consideradas tratamento de dados pessoais, bem como estendeu também a noção de dados pessoais, em relação às informações que identificam ou tornam identificável uma pessoa, incluindo dados genéticos, biométricos, fisiológicos, comportamentais, incluindo imagens faciais e dados datiloscópios (STANZIONE: 2016, p. 23/24). Com isso, a *privacy* viu-se transformada em direito de seguir as próprias informações onde quer que se encontrem e de opor-se às suas manipulações (DI GENIO: 2016, p. 164). Logo, verifica-se que na União Europeia a coleta, o tratamento e a utilização dos dados

deverão observar os seguintes princípios: a) da finalidade, ou seja, o dado fornecido pelo usuário não pode ser utilizado para outra situação senão o fim específico que gerou o fornecimento do mesmo; b) do consentimento, no sentido de que deve dar a expressa e específica a aquiescência; c) da minimização e proporcionalidade dos dados a recolher, a fim de evitar abusos; d) a supervisão e controle institucionais; e) o direitos dos titulares dos dados à informação e retificação (GONÇALVES, 2017). Também se pode acrescentar o princípio da responsabilidade, previsto no art. 82 do Regulamento, que dispõe que quem quer que sofra um dano patrimonial ou não patrimonial causado pela violação do Regulamento tem direito a obter a reparação de tais danos do responsável pelo tratamento dos dados (SICA: 2016, p. 8). Aliás, defende-se que "no Regulamento é mais amplo o âmbito da responsabilidade do titular do tratamento" (PARISI: 2016, p. 300).

Os especialistas europeus põem em relevo que há uma diferença sistêmica entre a antiga Diretiva de 1995 e o atual Regulamento de 2016. Em primeiro lugar, as diretivas da União Europeia fixam parâmetros a serem posteriormente adotados na legislação de cada país, que, todavia, conservam uma margem de discricionariedade na regulamentação interna. Já os regulamentos, por sua vez, têm aplicação direta em todos os países membros da U.E., dispensando e impedindo o acolhimento legislativo nacional. Uma tal opção revela a necessidade sentida de se regular de modo uniforme a questão em todo o espaço da União Europeia. Por outro lado, a forma como foi disciplinada a matéria revela que a proteção dos dados pessoais deixou de ser considerado apenas como um direito fundamental da pessoa, mas passou a ser uma questão de interesse público europeu, o que é deixado claro na redação do seu considerando 4, ao estabelecer que "o direito à proteção dos dados de caráter pessoal não é uma prerrogativa absoluta, devendo ser considerado à luz de sua função social, levando em consideração outros direitos fundamentais, observando-se o princípio da proporcionalidade". Mais claramente ainda, da combinação dos artigos 33 e 34 do Regulamento, percebe-se que a violação dos dados que comporte riscos para os direitos e liberdade das pessoas físicas deve sempre ser denunciada à Autoridade que exercita a função de controle (interesse público), sendo que tal violação somente será notificada ao interessado se tal risco for "elevado", ou se a referida Autoridade assim determinar (PIZZETTI: 2016, p. 4 e 5). É tendo estas experiências como ponto de partida, que se passa a tecer considerações acerca de critérios para a adequada implementação da mineração de dados, na ordem jurídica brasileira.

3.2.2 Critérios para a sua adequada implementação

A partir da análise do cenário normativo nacional, depreende-se que o Brasil está entre os países com maior deficiência, em matéria de proteção de dados, por lhe faltar uma autoridade oficial competente para a proteção de dados, bem como em face de inexistir lei específica sobre o tema. Contudo, não se está em um estado de completa anomia, já que os direitos à intimidade e à vida privada estão insculpidos na Carta Constitucional, bem como são previstos como direitos da personalidade em nosso Código Civil. Ademais, com o advento do Marco Civil da Internet (MCI), Lei sob o nº 12.965/14, dispositivos específicos trataram sobre a matéria, preservando a coleta, o tratamento e a utilização dos dados.

O artigo 3º do Marco Civil da Internet destaca expressamente o princípio de proteção de dados. O mesmo dispositivo refere "na forma da lei", como a apontar futura legislação que virá. Mencione-se, a respeito, o projeto de lei que tramita junto à Câmara dos Deputados, Projeto de Lei 5.276 de 2016, que dispõe sobre o "tratamento de dados pessoais para garantia do livre desenvolvimento da personalidade e da dignidade da pessoa natural" (BRASIL, 2016). Por sua vez, o artigo 7º, VII, assegura ao internauta o não fornecimento a terceiros de seus dados pessoais, inclusive registros de conexão e de acesso a aplicações de internet, salvo mediante consentimento livre, expresso e informado ou nas hipóteses previstas em lei. Portanto, como se vê, a transferência de dados para terceiro, somente poderá se operar, com consentimento expresso do autor, configurando-se a inobservância dessa restrição uma violação de direito à privacidade. Ademais, quanto à coleta e utilização dos dados, a legislação é clara no seguinte sentido:

> Art. 7o O acesso à internet é essencial ao exercício da cidadania, e ao usuário são assegurados os seguintes direitos: [...] VIII - informações claras e completas sobre coleta, uso, armazenamento, tratamento e proteção de seus dados pessoais, que somente poderão ser utilizados para finalidades que: a) justifiquem sua coleta; b) não sejam vedadas pela legislação; e c) estejam especificadas nos contratos de prestação de serviços ou em termos de uso de aplicações de internet;

Logo, mais uma vez, há determinação expressa acerca de que o dado não pode ser coletado, tratado, ou, ainda, objeto de uma análise preditiva, senão forem atendidas as finalidades previstas em lei e que os usuários tenham sido cientificados quando de sua coleta. Nesse sentido, ali estão os critérios básicos para mineração de dados. De igual sorte, cumpre colacionar o artigo 7º, em seu inciso IX, que consagra o consentimento expresso: "IX - consentimento expresso sobre coleta, uso, armazenamento e tratamento de dados pessoais, que

deverá ocorrer de forma destacada das demais cláusulas contratuais;". Dessa forma, tem-se que para a aplicação da mineração de dados, à luz do modelo europeu que: a) não podem ser utilizados dados para finalidades diversas daquelas para as quais foram fornecidos; c) deverá haver autorização expressa tanto para a coleta, como para o tratamento e utilização. Assim, deve haver aquiescência também para a análise preditiva; d) Não poderá haver abusos na coleta dos dados. Ademais, da análise perfunctória do Projeto de Lei (PL) 5.276 de 2016, depreende-se que, em seu artigo 6º, estão presentes restrições outras, além da finalidade e adequação, quais sejam, a necessidade, o livre acesso, a qualidade dos dados, a transparência, segurança, a prevenção e a não discriminação. Importa destacar, portanto, que tanto a coleta de dado sensível como a predição, que envolva *conhecimento sensível*, é matéria que pode gerar responsabilidade civil, por violação ao direito de privacidade.

4 CONSIDERAÇÕES FINAIS

A partir do estudo realizado, tornou-se possível tecer as seguintes considerações finais, a saber: 1) A expressão "mineração de dados" representa uma fonte geradora de conhecimento (output), a partir de dados fornecidos pelos próprios internautas (input), impactando diretamente a tomada de decisões pelos seus gestores; 2) Opera-se a análise preditiva quando se busca prever, antecipar situações, projetar eventos futuros, através de algoritmos computacionais, sem a interação humana; 3) São possíveis violações ao direito à privacidade pela prática da mineração de dados, em especial, pela ferramenta da análise preditiva. Como não se pode proibir tal prática, impõe-se o estabelecimento de restrições e salvaguardas normativas, a fim de se resguardar o mais possível o direito à privacidade; 4) É relevante o aprofundamento de critérios jurídicos para a aplicação de mineração de dados, na ordem jurídica brasileira. Dentre estes, destacam-se os princípios: a) da finalidade, ou seja, o dado que alcançado pelo usuário não pode ser utilizado para outra situação senão o fim específico que gerou o fornecimento pelo usuário; b) do consentimento, no sentido de que deve dar a expressa e específica aquiescência; c) da minimização e proporcionalidade dos dados a recolher, a fim de evitar abusos. 5) Tanto a coleta de dado sensível como a predição, que envolva "conhecimento" sensível, é matéria que pode gerar responsabilidade civil, por violação ao direito de privacidade.

Assim, as considerações supra procuram chamar a atenção para o fato de que, por

inafastável que seja o progresso tecnológico no mundo da informática, e por enormes que sejam as vantagens que ele traga para o mundo real em que vivemos, não se pode jamais perder de vista que a pessoa humana está e deverá permanecer no centro da ordem jurídica, cujos direitos fundamentais deverão ser protegidos até mesmo diante de necessidades e realidades do mercado.

REFERÊNCIAS

ASCENSÃO, José de Oliveira. **Estudos sobre Direito da Internet e da Sociedade da Informação.** Coimbra: Almedina, 2001.

AZAMBUJA, Celso Candido de. **Psiquismo digital:** sociedade, cultura e subjetividade na era da comunicação digital. Nova Petrópolis: Nova Harmonia, 2012.

BARROS, Augusto Paes de. **Trilhas em segurança da informação:** caminhos e ideias para a proteção de dados. Rio de Janeiro: Brasport, 2015.

BRASIL. **PL 5276/2016**. Disponível em: <http://www.camara.gov.br/proposicoesWeb/fichadetramitacao?idProposicao=2084378>. Acesso em: 2017.

CALIFANO, Licia. **Privacy: affermazione e pratica di um diritto fondamentale.** Napoli: Editoriale Scientifica, 2016.

CHERRY, Denny. **Fundamentos da privacidade digital.** São Paulo: Elsevier, 2014.

CREMER, Hans-Joachim. **Human Rights ant the Protection of Privacy in Tort Law –** A Comparison between English and German Law. New York: Routledge-Cavendish, 2011.

DI GENIO, Giuseppe. "Trasparenza e acesso ai datti personali". In: SICA, Salvatore; D'ANTONIO, Virgilio; RICCIO, Giovanni Maria (org.). **La nuova disciplina europea della privacy.** Milanofiori Assago: Wolters Kluwer Italia/CEDAM, 2016.

DONEDA, Danilo. **Da privacidade à proteção de dados pessoais.** Rio de Janeiro: Renovar, 2006.

GONÇALVES, Maria Eduarda. **A proteção de dados:** enquadramento. Disponível em: <http://opj.ces.uc.pt/e-learning/moodle/file.php/40/II_Notas_de_Enquadramento_Mo_dulo_V.pdf>. Acesso em: 2 ago. 2017.

HUNTER-HENIN, Myrian. "Horizontal Application and the Triumph of the European Convention on Human Rights". In: OLIVER, Dawn & FEDTKE, Jörg (ed.). **Human Rights and the Private Sphere –** A Comparative Study. Abingdon/UK: Routledge-Cavendish, 2007.

KRZEMINSKA-VAMVAKA, Joanna; O'CALLAGHAN, Patrick. "Mapping out a right of

privacy in tort law". In: BRÜGGEMEIER, Gert; CIACCHI, Aurelia Colombi; COMANDÉ, Giovanni. **Fundamental Rights and Private Law in the European Union –** Vol II: Comparative Analyses of Selected Case Patterns. Cambridge: Cambridge University Press, 2010.

MARQUESONE, Rosangela. **Big data:** técnicas e tecnologias para a extração de valor de dados. São Paulo: Casa do Código, 2016.

MAYER-SCHÖNBERGER, Viktor. **Delete –** The Virtue of Forgetting in the Digital Age. Princeton: Princeton University Press: 2009.

MENDES, Laura Schertel. **Privacidade, proteção de dados e defesa do consumidor.** São Paulo: Saraiva, 2014.

PAESANI, Liliana Minardi. **Direito e Internet –** Liberdade de informação, privacidade e responsabilidade civil. 5ª ed. São Paulo: Atlas, 2012.

PAGALLO, Ugo. **La tutela della privacy negli Stati Uniti D'America e in Europa:** modelli giuridici a confronto. Milano: Giuffrè, 2008.

_____. Sul Principio di Responsabilità Giuridica in Rete. **Il Diritto dell'Informazione e Dell'Informatica,** Roma, v. 25, n.4-5, p. 705-734, jul./out. 2009.

PALAZZI, Pablo A. "La intimidad de las telecomunicaciones a partir del fallo de la Corte Suprema en el caso 'Halabi'." In: POULLET, Yves; ASINARI, Maria Verônica Pérez; PALAZZI, Pablo (coord.). **Derecho a la intimidad y a la protección de datos personales.** Buenos Aires: Heliasta, 2009.

PARISI, Annamaria Giulia. "Responsabilità e sanzioni". In: SICA, Salvatore; D'ANTONIO, Virgilio; RICCIO, Giovanni Maria (org.). **La nuova disciplina europea della privacy.** Milanofiori Assago: Wolters Kluwer Italia/CEDAM, 2016.

PEREIRA, Ana Paula. **O que é um algoritmo?** Disponível em: <https://www.tecmundo.com.br/programacao/2082-o-que-e-algoritmo-.htm>. Acesso em: 2 ago. 2017.

PIZZETTI, Franco. **Privacy e il diritto europeo alla protezione dei dati personali** – II - Il Regolamento europeo 2016/679. Torino: Giappichelli, 2016.

POLÍTICA de dados. Disponível em: <https://www.facebook.com/about/privacy/>. Acesso em: 2017a.

POLÍTICA de privacidade. Disponível em: <https://www.google.com/intl/pt-BR/policies/privacy/>. Acesso em: 9 ago. 2017b.

POULLET, Yves; ASINARI, Maria Verônica Pérez; PALAZZI, Pablo (coord.). **Derecho a la intimidad y a la protección de datos personales.** Buenos Aires: Heliasta, 2009.

REHM, Gebhard M. "Privacy in the Digital Age: Vanishing in Cyberspace?" In: FRIEDMAN, Daniel; BARAK-EREZ, Daphne (ed.). **Human Rights in Private law.** Portland/Oregon: Hart

Publishing, 2003.

RODOTÀ, Stefano. **A vida na sociedade da vigilância –** A privacidade hoje. (Organização, seleção e apresentação de Maria Celina Bodin de Moraes). Rio de Janeiro: Renovar, 2008.

RODOTÀ, Stefano. **La rivoluzione della dignità**. Napoli: La Scuola di Pitagora, 2013.

SCHWAB, Klaus. **The fourth industrial revolution.** Genebra: World Economic Forum, 2016.

SICA, Salvatore. "Verso l'unificazione del diritto europeo alla tutela dei dati personali? In: SICA, Salvatore; D'ANTONIO, Virgilio; RICCIO, Giovanni Maria (org.). **La nuova disciplina europea della privacy.** Milanofiori Assago: Wolters Kluwer Italia/CEDAM, 2016.

SILVA, Leandro Augusto; PERES, Sarajane Marques; BOSCARIOLI. **Introdução à mineração de dados.** Rio de Janeiro: Elsevier, 2016.

SIMITIS, Spiros. "Reviewing Privacy in an Information Society". In: **University of Pennsylvania Law Review**, vol. 135 (1987).

SILVEIRA, Sérgio Amadeu; AVELINO, Rodolfo; SOUZA, Joyce. Sergio Amadeu da Silveira. A privacidade e o mercado de dados pessoais. **Liinc em Revista,** Rio de Janeiro, v. 12, n. 2, p. 217-230, nov. 2016. Disponível em: <http://www.ibict.br/liinc http://dx.doi.org/10.18617/liinc.v12i2.902>.

STANZIONE, Maria Gabriella. "Genesi ed ambito di applicazione". In: SICA, Salvatore; D'ANTONIO, Virgilio; RICCIO, Giovanni Maria (org.). **La nuova disciplina europea della privacy.** Milanofiori Assago: Wolters Kluwer Italia/CEDAM, 2016.

SOLOVE, Daniel J. "La persona digital y el futuro de la intimidad". In: POULLET, Yves; ASINARI, Maria Verônica Pérez; PALAZZI, Pablo (coord.). **Derecho a la intimidad y a la protección de datos personales.** Buenos Aires: Heliasta, 2009.