OS *BIG DATA* E OS DADOS PESSOAIS ENTRE OS PRINCÍPIOS DA PROTEÇÃO E DA INOVAÇÃO*

Big Data and Personal Data Between the Principles of Protection and Innovation

Submitted: 07/08/2019
Revised: 12/11/2019
Accepted: 22/01/2020

Célia Zolynski**

Abstract

Objective – *The article contrasts the problem of Big Data with the possibilities and limits of personal data protection. It is an original contribution to the academic discussion about the regulation of the Internet and the management of algorithms, focusing on Big Data.*

Methodology/approach/design – *The article provides bibliographic research on the opposition between Big Data and personal data protection, focusing on European Union law and French law. From the research is possible to identify regulatory alternatives do Big Data, whether legal-administrative nature or technological nature.*

Findings – *The article enlightens that, in addition to the traditional regulatory options, based on the law, there are technological options for regulating Big Data and algorithms. The article goes through an analysis of administrative performance, such as France's CNIL (*Commission nationale informatique et libertés*, CNIL), to show that it has limits. Thus, the article concludes that there is a need to build a new type of regulation, one that is open to the inputs of regulated parties and civil society, in the form of new co-regulatory arrangements.*

Practical implications – *The article has an obvious application since the production of legal solutions for Internet regulation requires combining them with technological*

*Texto derivado do seminário internacional "A efetividade do direito em face do poder dos gigantes da Internet – Brasil e França", realizado na Universidade de Brasília no período de 13 até 15 de abril de 2016. Agradece-se ao fomento da Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), da Fundação de Apoio à Pesquisa do Distrito Federal (FAPDF), da Embaixada da França no Brasil e das universidades brasileiras e francesas envolvidas. Tradução de Germana Henriques Pereira. Revisão técnica de Alexandre Veronese.

**Professora Titular (*Professeur Agregée*) na *Université Paris 1 Panthéon-Sorbonne* e do Institut de Recherche Juridiques de la Sorbonne (IRJS). Ela pesquisa na área de Direito digital, Direito da Propriedade Intelectual, Direito da União Europeia, Direito Comercial e liberdades fundamentais. Entre suas publicações destacam-se: ZOLYNSKI, Célia. *Méthode de transposition des directives communautaires: étude à partir de l'exemple du droit d'auteur et des droits voisins*. Paris: Dalloz, 2007; FAUVARQUE-COSSON, Bénédicte (dir.); ZOLYNSKI, Célia (dir.). *Le cloud computing – l'informatique en nuage*. Paris: Société de Législation Comparé, 2014; MARTIAL-BRAZ, Nathalie; ZOLYNSKI, Célia. *La gratuité: un concept aux frontières de l'économie et le droit*. Paris: LGDJ / Lextenso Éditions, 2013. E-mail: celia.zolynski@univ-paris1.fr.

solutions. Brazil and several Latin American countries are experiencing this agenda, as they are building institutions and solutions to solve the dilemma of personal data protection.

Originality/value – *The article clarifies several parts of the General Data Protection Regulation (EU Regulation 2016/679) and its applicability to Big Data. These new types of data processing impose several legal and regulatory challenges, whose solutions cannot be trivial and will rely on new theories and practices*.

Keywords: *Big Data. Algorithm regulation. Personal data protection. European Union Law. French Law.*

Resumo

Objetivo – O artigo contrapõe o problema dos Big Data com as possibilidades e limites da proteção de dados pessoais. Ele é uma contribuição original ao sobre regulação da Internet e da gestão de algoritmos, com foco nos Big Data.

Metodologia – O artigo comunica uma pesquisa bibliográfica sobre o tema da contraposição dos *Big Data* e da proteção de dados pessoais, com foco no direito da União Europeia e no direito francês. A partir da pesquisa bibliográfica, foi possível identificar alternativas regulatórias, de caráter jurídico-administrativo e de caráter tecnológico.

Resultados – O artigo demonstra que, além das tradicionais opções regulatórias, com base no direito, existem opções tecnológicas para regulação dos *Big Data* e dos algoritmos. O artigo percorre uma análise da atuação administrativa – autoridades de dados pessoais, como a Comissão Nacional Informática e Liberdades (*Commission Nationale Informatique et Libertés* – CNIL) da França – para mostrar que ela possui limites. Assim, o artigo conclui que há a necessidade de se construir um novo tipo de regulação, que seja permeável aos influxos dos regulados e da sociedade civil, na forma de arranjos corregulatórios.

Implicações práticas – O artigo possui evidente aplicação, uma vez que a produção de soluções jurídicas, para regulação da Internet, exige o seu conjugar com soluções de caráter tecnológico. O Brasil e diversos países da América Latina estão sintonizados nesse tema, uma vez que estão a construir instituições e soluções para resolver o dilema para proteção de dados pessoais.

Originalidade/relevância do texto – O texto esclarece diversos trechos do Regulamento Geral sobre a Proteção de Dados Pessoais (Regulamento UE 2016/679) e sua aplicabilidade em relação aos Big Data. Esses novos tipos de tratamentos de dados pessoais impõem diversos desafios jurídicos e regulatórios, cujas soluções não poderão ser triviais e dependerão de novas teorias e práticas.

Palavras-chave: *Big Data*. Regulação de algoritmos. Proteção de dados pessoais. Direito da União Europeia. Direito da França.

INTRODUÇÃO: OS *BIG DATA*, UMA REVOLUÇÃO DOS DADOS

Os *Big Data*[1]: esse termo encarna uma revolução. A revolução dos dados. Uma revolução da informação comparável à primeira revolução industrial (MAYER-SCHÖNBERGER e CUKIER, 2014). Incontestavelmente, os *Big Data* marcam, no mínimo, a nossa entrada na fase mais atual da Sociedade da Informação. O termo *Big Data* refere-se, mais precisamente, à capacidade de tratar[2], em tempo quase real, enormes repositórios de informação, grandes quantidades de dados estruturados, semiestruturados ou até mesmo não estruturados e díspares. A partir desse tratamento é possível encontrar novos dados que, sem tal processo, ficariam desconhecidos. Os *Big Data* são definidos, na realidade, por referência à regra dos "5V" (FRANÇA, 2014). O primeiro "v" refere-se a um grande *volume* de dados, ao qual se acrescenta uma grande *variedade*, uma vez que os *Big Data* permitem a recolha de dados heterogêneos, em vários formatos, estruturados ou não estruturados. Esses dois primeiros "v" são incentivados pela explosão de dados produzidos pelas empresas, pelo Estado ou pelos indivíduos em um movimento generalizado de recolha de dados. Esses dados podem estar estruturados em bases, eles podem advir de intercâmbios em redes sociais, eles podem ser dados públicos abertos, eles podem derivar de dados de navegação ou, ainda, eles podem ser dados recolhidos a partir de objetos conectados, dentre outras. O terceiro "v" refere-se à *velocidade* no tratamento dos dados, na medida em que ela pode, agora, aumentar até quase atingir o tempo real. Atualmente, já é possível capturar uma massa de dados, derivada de dispositivos móveis, para correlacioná-la dinamicamente. O aumento na capacidade de tratamento ocorreu por processos tecnológicos atuais, tais como o *outsourcing*[3] proporcionado pela computação em nuvem. Também, há que mencionar o poder cada vez maior do tratamento algorítmico desses dados, o qual permite que qualquer empresa tenha acesso à tecnologia de *Big Data*. Acrescenta-se aos três "v", ainda, o quarto, que se refere à *veracidade*, a qual sublinha o desafio relativo à fiabilidade dos dados explorados. E, importa destacar, acima de tudo, o quinto "v", de *valor*. Afinal, os *Big Data* geram um novo *valor* ao transformar os dados em informações úteis, de uma maneira inovadora. Os novos usos combinados de dados e de novas ferramentas de compreensão estão se desenvolvendo. E eles

[1] Nota do revisor técnico. Cabe frisar que a palavra *data* está no plural e em latim, cujo singular é *datum* e se refere a qualquer informação em seu estado natural, ou seja, antes de um tratamento (BUCKLAND, 2017).

[2] Nota do revisor técnico. Optou-se pela utilização do verbo tratar na tradução, uma vez que ele é utilizado na legislação brasileira e no direito da União Europeia, na sua versão oficial em português.

[3] Nota do revisor técnico. O termo *outsourcing* é um tipo específico de terceirização, usado nas áreas de tecnologia da informação e de informática, que se refere ao tratamento externo de dados de uma organização. Como as organizações externas são mais especializadas, elas podem ofertar esses serviços de forma mais eficiente.

estão revolucionando os métodos de análise disponíveis. Os *Big Data* inscrevem-se numa inversão do método estatístico clássico que procede de uma lógica dedutiva cujas conclusões se baseiam numa amostragem através de causalidades. Eles repousam, ao contrário, numa lógica indutiva que leva a coletar o maior número possível de dados para explorar as suas correlações. Essa massa de dados é, então, misturada para se extrair, através dessas correlações, novas informações. Tais novas informações podem ser entendidas como "sinais fracos", que são muito úteis para análises que são, frequentemente, preditivas. Os dados, cujo valor latente é revelado nos *Big Data*, irão, então, quase que "falar por si", assim como serão capazes de fornecer uma resposta sem que uma pergunta prévia sequer tenha sido colocada (MAYER-SCHÖNBERGER e CUKIER, 2014, p. 121 *et seq*.).

Quais são os desafios propostos pelos *Big Data*? Eles oferecem oportunidades sem precedentes para operadores econômicos, como empresas, e, também, para as autoridades públicas. A fusão de dados e a sua análise preditiva são um bom presságio para um número considerável de novas aplicações. Dentre elas, é possível indicar a análise dos sentimentos, da segmentação e da geolocalização das necessidades das pessoas, gerando um conhecimento aprofundado do comportamento dos usuários de sistemas informatizados e, até mesmo, das suas expectativas para, então, adaptar as ofertas em tempo real (COINTOT e EYCHENNE, 2014, p. 31 *et seq*.). As externalidades variam desde a comercialização inteligente até à otimização dos processos de produção ou, ainda, até uma melhor gestão das cidades inteligentes (gestão da energia, do tráfego e até mesmo da segurança e da saúde pública), sem mencionar as perspectivas em termos de pesquisa científica básica ou fundamental (ESTADOS UNIDOS DA AMÉRICA, 2014). Mas, ao mesmo tempo, os *Big Data* são fontes de grandes riscos. Alguns denunciam as consequências negativas que esse poder de informação pode gerar se fosse reservado a alguns operadores dominantes (BENSAMOUN e ZOLYNSKI, 2015). Outros argumentam que existe um risco de censura ou de discriminação que as pessoas poderiam correr devido ao tratamento algorítmico dos seus dados, que teria por objetivo lhes oferecer serviços personalizados com preços dinâmicos que podem privá-las da sua liberdade de escolha (ZOLYNSKI, 2015). Sem mencionar o medo, que pode ser uma fantasia, de uma ditadura de dados e de algoritmos preditivos impondo uma definição de "ser humano calculado" que desafiaria o princípio da autodeterminação individual que, segundo a visão moderna, é a própria essência de toda pessoa humana (MAYER-SCHÖNBERGER e CUKIER, 2014, p. 185 *et seq*.).

Sendo assim, como pode a nossa sociedade tirar proveito útil dos *Big Data* enquanto gerencia o risco informacional associado com tais novos usos? Os operadores devem ser autorizados a tirar partido da massa de dados que dispõem

acerca dos seus clientes para valorizá-los? (WILLART e CRIÉ, 2017) Assim, por exemplo, poderiam proceder os bancos, os operadores de seguros ou as empresas telecomunicações, na medida em que essas operações com *Big Data* se tornariam diferenciais competitivos? Como é possível desenvolver serviços inovadores e competir com os principais operadores, em particular com os gigantes da Internet norte-americanos, cujo modelo de negócio se baseia na garimpagem maciça de dados? (SMYRNAIOS, 2017). Ao mesmo tempo, como valorizar essas utilizações, preservando simultaneamente os direitos de terceiros, tendo em conta os elevados riscos que os grandes volumes de dados podem gerar? Em especial, como evitar os grandes riscos de violação dos dados pessoais, da privacidade dos indivíduos e mesmo do seu direito à autodeterminação? Como será possível pensar esse novo equilíbrio?

Estes riscos relacionados com as informações extraídas dos *Big Data* são muito reais. Por conseguinte, é atualmente necessário aperfeiçoar a análise para encontrar o equilíbrio certo entre os perigos prováveis e o potencial de inovação que os *Big Data* suscitam. As apostas são altas e a análise delicada, especialmente porque o contexto é altamente evolutivo em termos técnicos e sociais: os *Big Data* representam um verdadeiro desafio a esse respeito. Coloca-se, pois, a questão de saber como conciliar a proteção dos indivíduos e a promoção da inovação, uma vez que as duas não são necessariamente mutuamente excludentes. A inovação e a proteção seriam justamente os dois princípios que condicionariam o desenvolvimento sustentável dessas novas utilizações e, de um modo mais geral, da economia baseada nos dados. Este é, de fato, o sentido das reformas que estão sendo pensadas atualmente, pelo menos na Europa, de acordo com tendências semelhantes, particularmente no Canadá, sob a liderança de Ann Cavoukian[4]. Portanto, parece necessário identificar quais soluções devem ser recomendadas para garantir o desenvolvimento de práticas de "dados responsáveis" para aproveitar as oportunidades sem precedentes dos *Big Data*. Para identificar essas soluções, primeiro é necessário rever as dificuldades colocadas pelos *Big Data* no contexto do direito de proteção dos dados pessoais, que serão analisadas na segunda seção. Depois, há que se considerar as formas de melhor gerir os riscos informacionais inerente aos *Big* Data, na terceira seção. A conclusão do presente trabalho aponta para a necessidade de construção de uma política de corregulação, que possa equilibrar os princípios indicados – inovação e proteção – com atenção aos dilemas do presente e do futuro. Ela enfatiza, ainda, a necessidade de que haja a participação da sociedade civil nesse debate, uma vez que os riscos existentes nos *Big Data* dizem respeito a todos.

[4] Nota do revisor técnico. Ela foi a Comissária para Informação e Privacidade da Província (Estado) de Ontário, do Canadá, entre 1997-2014. Ela se tornou conhecida por ter cunhado o conceito de "privacy by design" (RALLET, ROCHELANDET e ZOLYSNKI, 2015).

OS *BIG DATA* E OS DADOS PESSOAIS: QUAIS AS DIFICULDADES?

A primeira dificuldade a ser indicada está relacionada com a definição de dados pessoais no sentido jurídico do termo, em relação aos dados usados nos *Big Data*. Essa é uma questão fundamental, na medida em que condiciona a aplicação das normas de proteção dos indivíduos neste domínio. Como estabelecer uma relação clara entre os *Big Data* e os dados pessoais. Os grandes volumes de dados revelam, em especial, as dificuldades crescentes que podem surgir quando se torna necessário definir de modo preciso os dados pessoais dos indivíduos. Isto é particularmente evidente nos debates desencadeados pela construção do Regulamento Geral sobre a Proteção de Dados Pessoais (RGPD), aprovado em 2016 (UNIÃO EUROPEIA, 2016). Essas dificuldades são acrescidas, ao se tentar compreender as características específicas dos *Big Data*, ou seja, a sua natureza dinâmica, móvel, evolutiva e preditiva.

Há que reconhecer a inadequação da análise tradicional. Assim, é possível se interrogar a respeito dos dados capturados pelos *Big Data*: tratam-se de dados pessoais? Se a variedade dos *Big Data* leva a considerar que todos os tipos de dados têm valor, mesmo que sejam fragmentos de informação, então, uma grande quantidade dos dados recolhidos não será, em si mesma, rotulável como dados pessoais na acepção do direito positivo. O *considerandum* 26 do RGPD entabula esse debate sobre a possibilidade de identificação dos indivíduos:

> Os princípios da proteção de dados deverão aplicar-se a qualquer informação relativa a uma pessoa singular identificada ou identificável. Os dados pessoais que tenham sido pseudonimizados, que possam ser atribuídos a uma pessoa singular mediante a utilização de informações suplementares, deverão ser considerados informações sobre uma pessoa singular identificável. Para determinar se uma pessoa singular é identificável, importa considerar todos os meios suscetíveis de ser razoavelmente utilizados, tais como a seleção, quer pelo responsável pelo tratamento quer por outra pessoa, para identificar direta ou indiretamente a pessoa singular. Para determinar se há uma probabilidade razoável de os meios serem utilizados para identificar a pessoa singular, importa considerar todos os fatores objetivos, como os custos e o tempo necessário para a identificação, tendo em conta a tecnologia disponível à data do tratamento dos dados e a evolução tecnológica. Os princípios da proteção de dados não deverão, pois, aplicar-se às informações anónimas, ou seja, às informações que não digam respeito a uma pessoa singular identificada ou identificável nem a dados pessoais tornados de tal modo anónimos que o seu titular não seja ou já não possa ser identificado. O presente regulamento não diz, por isso, respeito ao

> tratamento dessas informações anónimas, inclusive para fins estatísticos ou de investigação (UNIÃO EUROPEIA, 2016).

O *considerandum* deixa o debate jurídico claro. Sobretudo porque a finalidade do tratamento de grandes volumes de dados – *Big Data* – não é, necessariamente, identificar ou fazer identificar uma pessoa, mas, na melhor das hipóteses, um comportamento. Basta pensar no processamento para fins estatísticos. A natureza coletiva do *Big Data* aumenta a dificuldade, uma vez que os dados não emanam, necessariamente, de um indivíduo – ou conjunto de indivíduos – definido; eles derivam da recolha de grandes massas de dados, do cruzamento de referências e de processos de dedução. No entanto, a natureza dinâmica dos *Big Data*, bem como as correlações e as interpretações, pode levar *a posteriori* a resultados que identifiquem os indivíduos. Para desvendar isso, seriam necessárias, portanto, análises diacrônicas, ou seja, que ocorram ao longo do ciclo de vida do tratamento para que se possa determinar a natureza pessoal, ou não, dos dados usados. Adotando-se uma abordagem alargada, seria possível até considerar, como alguns fazem, que qualquer tipo de tratamento de grandes volumes de dados poderia tornar a pessoa identificada ou identificável, especialmente em tempos de políticas de dados abertos e, por conseguinte, isso poderia implicar na qualificação de tais dados como pessoais (ROUVORY, 2014). Em reação a essa abordagem extensiva, outros consideram, pelo contrário, que o tratamento de grandes volumes de dados deveria, em princípio, levar à exclusão em si da qualificação dos dados pessoais. É isto que algumas empresas e juristas norte-americanos defendem, com o objetivo de não restringir o potencial de inovação dos grandes volumes de dados, sujeitando-os à proteção vinculativa dos dados pessoais (LATREILLE e ZOLYNSKI, 2014).

Uma solução possível seria adotar uma nova abordagem casuística. Entre estas duas concepções opostas, a diversidade do tratamento dos *Big Data* pressupõe certamente que não seja dada nenhuma resposta, em princípio, categórica, tal como "tudo é pessoal" ou "nada é pessoal". A análise deve ser feita caso a caso: trata-se de verificar se um dado inicialmente recolhido é pessoal; ou, se o resultado do tratamento dos *Big Data* tem por objetivo, ou por efeito, identificar ou, tornar uma pessoa identificável. No entanto, esse processo casuístico será, por vezes, muito delicado e a situação criará incerteza para se determinar se o direito de proteção de dados pessoais é aplicável, ou não. Nesse caso, portanto, essa insegurança não é capaz de assegurar uma proteção eficaz dos indivíduos, em equilíbrio com a promoção do potencial de inovação dos operadores. O casuísmo, em síntese, não seria a melhor solução possível.

Outra solução possível seria uma abordagem baseada no uso dos *Big Data*. Talvez fosse melhor modificar a abordagem. Assim, seria possível alterar o ângulo de análise com o objetivo de focalizar não somente os dados; mas, também

e principalmente, o uso que deles é feito. De fato, seriam os resultados e usos dos tratamentos dos *Big Data* que deveriam ser levados em consideração. Essa alteração de abordagem permitiria ir além da visão tradicional, que é historicamente estática, em relação aos dados e que é adotada pelo direito da proteção de dados pessoais, focalizados nos processos de recolha. Por conseguinte, seria adequado estabelecer uma nova abordagem, dinâmica e coletiva, que refletiria a utilização dos *Big Data*. É este o sentido das propostas apresentadas pelo Conselho de Estado francês no seu relatório de 2014, intitulado "O mundo digital e os direitos fundamentais". O relatório distingue assim três categorias diferentes de utilizações, que se referem ao termo genérico *Big Data*: uma primeira categoria refere-se aos usos que não remetem aos dados pessoais, como "a otimização da manutenção dos sistemas [veículos] responsáveis pela utilização de dados transmitidos por captadores" (FRANÇA, 2014, p. 172-173). Uma segunda categoria seria composta por usos relacionados com dados pessoais, mas cuja finalidade seria estatística, como, por exemplo, "a utilização de bases de dados de seguros de saúde para detectar reações adversas a um medicamento". A terceira categoria, por fim, compreenderia os usos não estatísticos de dados pessoais, como a publicidade comportamental. Para além disso, os usos resultantes dos tratamentos algorítmicos deveriam receber uma atenção especial, em virtude de seu impacto potencialmente decisivo sobre os direitos e as liberdades fundamentais. E, em especial, porque esses tratamentos poderão determinar o que é relevante, ou não, para o indivíduo, com um risco de erro gerado por um tratamento puramente racional dos comportamentos; um tratamento que nega qualquer dimensão de sensibilidades e, por vezes, inadequado para identificar as necessidades ou os comportamentos futuros. Consequentemente, parece que o enquadramento jurídico dos *Big Data* deveria consistir em interpretar nitidamente os diferentes usos, identificando aqueles que podem se mostrar perigosos aos indivíduos que o direito da proteção de dados visa proteger. A análise de parte da doutrina norte-americana contempla plenamente essa ideia, na medida em que leva ao reconhecimento de que os *Big Data* não seriam dados pessoais, uma vez que seriam anonimizados[5] e não causariam danos aos titulares. Aliás, a própria noção de dano está assim integrada na definição de dados da legislação canadense. A subseção 3 e os parágrafos (a) e (b) da seção 10 [Violação de Garantias de Segurança] da Lei sobre a proteção das informações pessoais e documentos eletrônicos:

[5] Cabe notar que o Regulamento Geral sobre a Proteção de Dados (RGPD) usa o termo "pseudonimização", ao passo em que a Lei nº 13.709/2018 usa o termo "anonimização". O termo, no português de Portugal, é mais próximo do francês: "pseudonymisation".

> (3) Além das circunstâncias listadas na subseção 7(3) [divulgação de dados pessoais do titular sem o seu conhecimento ou consentimento], para o objetivo da cláusula 4.3 do Anexo 1, apesar da ressalva que acompanha tal cláusula, uma organização pode divulgar informações pessoais sem o conhecimento ou consentimento do indivíduo se (a) tal divulgação é feita para outra organização, para instituição governamental, ou parte de uma instituição governamental que foi notificada da violação descrita na subseção (1); e (b) a divulgação é feita somente para as finalidades de reduzir os riscos de danos ao indivíduo que pode advir da violação, ou, para mitigar tal risco (CANADÁ, 2019).

No RGPD, a União Europeia procurou encontrar um equilíbrio entre a proteção e a inovação. Foram mantidos como parte dos princípios do direito da proteção de dados pessoais, os princípios da proporcionalidade[6], da limitação da conservação (artigo 5º, parágrafo 1, "e" do RGPD), do consentimento prévio[7] e, sobretudo, da limitação da finalidade (artigo 5º, parágrafo 1, "b" do RGPD), apesar de alguns autores considerarem que eles teriam se tornado inadequados para a natureza dinâmica, mutável, evolutiva e preditiva dos *Big Data* (BENSAMOUN e ZOLYNSKI, 2015). No entanto, esses princípios são flexibilizados para permitir a compreensão das especificidades dessas novas técnicas de valorização dos dados. Por exemplo, o princípio da finalidade é flexibilizado para permitir uma "finalidade compatível" com a finalidade da coleta original (MATTATIA, 2016). Além disso, o RGPD pretende facilitar o acesso aos *Big Data* para fins de investigação no domínio da saúde pública, como alguns recomendam atualmente. Ou, ainda, ele visa também promover as operações de tratamento estatístico consideradas promissores, submetendo-as a uma presunção de legalidade quando estejam sujeitas a medidas adequadas de anonimização. Além disso, em sentido contrário, o *datamarketing* e, de um modo mais amplo, os tratamentos destinados a identificar o indivíduo ou o seu comportamento econômico são estritamente regulamentados pelo RGPD. Desse modo, os tratamentos algorítmicos serão rigorosamente regulamentados quando constituírem fonte de discriminação negativa, ainda que fosse preferível que RGPD fosse mais rígido no âmbito dos usos dos tratamentos utilizações algorítmicas, impondo aos operadores uma obrigação de lealdade. Afinal, tais tratamento dos *Big Data* requerem a delimitação precisa dos riscos informacionais

[6] Nota do revisor técnico. Esse princípio geral está bem tratado nos *consideranda* 4, 19, 49, 50, 62, 73, 129, 151, 152, 156 e 170. Ainda, a proporcionalidade de aplicação do RGPD é mencionada em diversos dispositivos normativos, também. Ao se tratar da licitude dos tratamentos, no RGPD se indica que a finalidade do mesmo deve ser "proporcional ao objetivo legítimo prosseguido" (artigo 6º, parágrafo 3, "b"). Também, a proporcionalidade deve ser observada na avaliação de impacto (artigo 35º, parágrafo 7, "b"), bem como dentre outros procedimentos, como aplicação das sanções.

[7] Nota do revisor técnico. Previsto em vários *consideranda*, bem como nos artigos 6º, 7º, 8º, e 9º. O consentimento é um dos modos de garantir licitude para um tratamento de dados pessoais.

reais e potenciais. Uma vez identificados, convém considerar os meios a ser mobilizados para se assegurar uma gestão adequada desses riscos informacionais nos *Big Data*. Esse será o tema central da próxima seção, que analisará as possibilidades de gerenciamento dos riscos informacionais.

INSTRUMENTOS PARA GERIR O RISCO INFORMACIONAL DOS *BIG DATA* E O PAPEL DO FORTALECIMENTO DOS TITULARES DOS DADOS PESSOAIS

A gestão do risco informacional resultante dos *Big Data* exige, em primeiro lugar, que se definam instrumentos jurídicos destinados a responsabilizar os responsáveis pelos tratamentos, como será mostrado na próxima subseção. Mas, somente isso não será suficiente. É igualmente necessário assegurar proteção aos titulares dos dados pessoais, atribuindo-lhes um papel ativo no controle desses usos, como será descrito na segunda subseção.

Responsabilizar os Responsáveis pelos Tratamentos

O Direito da União Europeia e as autoridades de proteção de dados atualmente pretendem promover uma política de responsabilização dos responsáveis pelos tratamentos que desejam tirar partido do potencial de inovação dos *Big Data*. Essa prática de *accountability* remete à abordagem norte-americana e exige que os operadores tanto avaliem, quanto protejam os titulares dos dados pessoais danos efetivos e potenciais causados pelos *Big Data*. Tal responsabilização parece ser justificável em virtude dos seus conhecimentos acerca dos usos feitos a partir dos tratamentos e dos benefícios que eles pretendem obter dos *Big Data*. Vários instrumentos jurídicos foram desenhados para cumprir as obrigações de licitude, transparência e de lealdade, as quais os operadores estão sujeitos no contexto do tratamento de grandes volumes de dados.

As obrigações de Licitude, de Transparência, de Lealdade e de Salvaguarda Contínua

Os responsáveis pelos tratamentos estão assim sujeitos às obrigações de licitude, de transparência, de lealdade e de salvaguarda contínua[8], as quais estão na base da análise de risco que agora são obrigados a realizar. O RGPD considera que a análise de risco é o método a utilizar para garantir a proteção dos dados pessoais, preservando simultaneamente o potencial de inovação das empresas. Em

[8] Nota do revisor técnico. O texto original usava o termo *vigilance*. Contudo, a tradução literal para vigilância poderia induzir confusão, uma vez que o objetivo da comunicação era demonstrar o dever de zelar continuamente e não de vigiar ou – até - espionar. Porém, o termo zelo também não seria a melhor expressão. A expressão salvaguarda contínua traduz melhor a ideia original.

primeiro lugar, os responsáveis pelo tratamento estão sujeitos a um dever de salvaguarda continuado. Eles são obrigados a realizar uma análise de risco antes e durante todo o ciclo de tratamento de grandes volumes de dados no que respeita a potenciais violações de dados pessoais, como disposto no parágrafo 1 do artigo 35º:

> Quando um certo tipo de tratamento, em particular que utilize novas tecnologias e tendo em conta a sua natureza, âmbito, contexto e finalidades, for suscetível de implicar um elevado risco para os direitos e liberdades das pessoas singulares, o responsável pelo tratamento procede, antes de iniciar o tratamento, a uma avaliação de impacto das operações de tratamento previstas sobre a proteção de dados pessoais (UNIÃO EUROPEIA, 2016).

No caso da França, para as operações de tratamento mais arriscadas, os responsáveis deverão realizar uma análise de impacto na privacidade (*privacy impact assessment*, ou PIA) (FRANÇA, 2018, p. 6). Trata-se, em primeiro lugar, de avaliar o contexto para se estimar o nível de riscos ou de ameaças, dependendo: da magnitude dos riscos, de sua gravidade; e, de sua verossimilhança, ou seja, de sua probabilidade de ocorrência. Nesse contexto, é necessário prever as medidas necessárias e suficientes para gerir esse risco e respeitar as obrigações jurídicas relativas ao tratamento dos dados. Algumas especificações podem ser fornecidas acerca desse método de gestão dos riscos informacionais. Convém notar que a importância destas obrigações depende do grau do risco gerado pelo tratamento. Assim, trata-se de respeitar o mencionado princípio de proporcionalidade, que se traduz na obrigação de realizar estudos de impacto, quando for o caso. Depois, é essencial pensar nessas medidas ao longo do tempo. Em outras palavras, a análise de risco deve ser realizada ao longo de todo o projeto para que as medidas sejam eventualmente adaptadas para levar em conta, por exemplo, a evolução dos tratamentos algorítmicos ou o cruzamento de novas massas de dados; é por esse motivo que se menciona a salvaguarda contínua. O objetivo central será responder aos riscos de perda da anonimização, ligados ao desenvolvimento de novas técnicas algorítmicas ou de referências cruzadas com novas bases de dados. Por conseguinte, no que diz respeito aos *Big Data*, recomenda-se particularmente a adoção de uma abordagem de *Privacy by Design*. Trata-se de uma abordagem cuja origem está na proposição de Ann Cavoukian[9], a qual, apesar de não ter sido integrada no direito canadense, atualmente, está consagrada no artigo 25º do RGPD como um ponto fundamental de proteção dos dados pessoais:

[9] Nota do revisor técnico. Já explicado na nota 14. Releva mencionar que a Câmara dos Comuns do Canadá produziu um relatório acerca da necessidade de revisar a legislação nacional canadense para se adequar ao RGPD, com destaque para a questão do *Privacy by Design* (CANADÁ, 2018).

> Artigo 25º. - Proteção de dados desde a conceção e por defeito[10]. 1. Tendo em conta as técnicas mais avançadas, os custos da sua aplicação, e a natureza, o âmbito, o contexto e as finalidades do tratamento dos dados, bem como os riscos decorrentes do tratamento para os direitos e liberdades das pessoas singulares, cuja probabilidade e gravidade podem ser variáveis, o responsável pelo tratamento aplica, tanto no momento de definição dos meios de tratamento como no momento do próprio tratamento, as medidas técnicas e organizativas adequadas, como a pseudonimização, destinadas a aplicar com eficácia os princípios da proteção de dados, tais como a minimização, e a incluir as garantias necessárias no tratamento, de uma forma que este cumpra os requisitos do presente regulamento e proteja os direitos dos titulares dos dados. 2. O responsável pelo tratamento aplica medidas técnicas e organizativas para assegurar que, por defeito, só sejam tratados os dados pessoais que forem necessários para cada finalidade específica do tratamento. Essa obrigação aplica-se à quantidade de dados pessoais recolhidos, à extensão do seu tratamento, ao seu prazo de conservação e à sua acessibilidade. Em especial, essas medidas asseguram que, por defeito, os dados pessoais não sejam disponibilizados sem intervenção humana a um número indeterminado de pessoas singulares. 3. Pode ser utilizado como elemento para demonstrar o cumprimento das obrigações estabelecidas nos nºs. 1 e 2 do presente artigo, um procedimento de certificação aprovado nos termos do artigo 42º (UNIÃO EUROPEIA, 2016).

Segundo a compreensão o legislador da União Europeia, essa solução permitiria equilibrar os interesses entre a proteção e a inovação. De fato, com isso, fica consagrado um princípio de prevenção, o qual remete às abordagens – já adotadas em direito sanitário ou em direito ambiental – para gerir os riscos através da medição dos seus custos e dos benefícios esperados. Consequentemente, os responsáveis pelos tratamentos deverão, portanto, adotar medidas técnicas e organizativas para prevenir os riscos, em conformidade com essa nova abordagem proativa, ou seja, *ex ante* (ZOLYNSKI, 2016). As obrigações de salvaguarda contínua e de responsabilização são ampliadas por meio de uma obrigação de transparência, especialmente no tratamento de dados sensíveis. Isso se reflete no dever de manter a documentação dos processos, a qual incumbe aos responsáveis pelos tratamentos, como uma extensão da noção de auditoria (GAUTRAIS, 2014). Essa medida destina-se não somente a informar terceiros sobre as medidas aplicadas, mas também a permitir que as autoridades reguladoras exerçam controles de conformidades. Ela está prevista, como princípio, no artigo 5º, parágrafo 1, "a", e, como norma jurídica, no artigo 30º, do RGPD. Por conseguinte, os responsáveis pelos tratamentos terão que comunicar as medidas e garantias operacionais aplicadas para evitar ou remediar os riscos em caso de

[10] Cabe frisar que o "por defeito" da expressão do português de Portugal tem o sentido da expressão latina *per default*, que se traduz como "por padrão". É comum que, em inglês, se utilize *by default*. A expressão francesa é um latinismo: *par défaut*.

ocorrência nefastas (por exemplo, em caso de violação da segurança). Ainda, relatórios poderão ser exigidos dos responsáveis pelos tratamentos. Tal obrigação de dar transparência aos processos, poderá se assemelhar, no futuro, às obrigações – atuais – de publicação que sujeitam empresas com ações nas bolsas de valores em termos de responsabilidade social e ambiental. Também, cabe destacar que os responsáveis pelos tratamentos estão sujeitos a um princípio e uma obrigação de lealdade durante o processo de tratamento dos dados pessoais, como inscrito do artigo 5º, parágrafo, "a" do RGPD. Além disso, seria importante estender essa obrigação de lealdade à utilização de instrumentos dos *Big Data*, especificamente aqueles referidos à implementação do tratamento algorítmico, conforme recomendado pelo mencionado relatório do Conselho de Estado francês (FRANÇA, 2014). Ao se considerar a abertura do conceito de lealdade, seria muito relevante ampliar tal princípio para abarcar uma obrigação de vedação às discriminações, em especial no que diz respeito aos tratamentos algorítmicos.

Instrumentos para Promover esses Princípios e Obrigações

Atualmente estão sendo analisados diversos instrumentos jurídicos que visam esclarecer o alcance dessas obrigações de salvaguarda contínua, de transparência e de lealdade. Todavia, devem ser destacados tipos gerais. O primeiro se refere às ferramentas de regulação setorial, as quais permitem amoldar cartilhas de práticas que devem ser observadas por segmentos específicos. Dentre elas estão os códigos de conduta e os pacotes setoriais de conformidade. O segundo tipo geral é composto pelas cartas éticas.

Contemporaneamente, estão sendo desenvolvidos instrumentos reguladores na forma de instrumentos de *soft law* (BENSAMOUN e ZOLYSNKI, 2015). As autoridades da União Europeia incentivam a utilização da corregulação, que envolve os responsáveis pelos tratamentos e os órgãos reguladores no sentido de desenvolverem normas flexíveis, adaptadas às necessidades das práticas setoriais e adaptáveis às evoluções técnicas. Essas normas poderiam, por exemplo, assumir a forma de códigos de conduta, tal como recomendado pelo RGPD, no artigo 40º:

> Artigo 40º. Códigos de conduta.
>
> 1. Os Estados-Membros, as autoridades de controlo, o Comité e a Comissão promovem a elaboração de códigos de conduta destinados a contribuir para a correta aplicação do presente regulamento, tendo em conta as características dos diferentes setores de tratamento e as necessidades específicas das micro, pequenas e médias empresas (UNIÃO EUROPEIA, 2016).

Esses códigos estão sendo atualmente ser desenvolvidos pela francesa *Comission nationale informatique et libertés* (Comissão Nacional de Informática

e Liberdades, CNIL) com atores de diferentes setores para supervisionar as utilizações de grandes volumes de dados (FRANÇA, 2019). Essa abordagem "iterativa"[11] está agora sendo construída em alguns setores (TIC SANTÉ, 2019). Além deles, existem os pacotes de conformidade (*packs de conformité*), da CNIL (FRANÇA, 2020). Assim, já foram criados e adotado pacotes de conformidade em vários setores: medidores de consumo (*compteurs communicants;* AMR – *automated meter reading*); no setor de alojamentos sociais (políticas de moradia para pessoas de baixa renda); seguros; veículos automatizados; e, da indústria de prata (setores de produtos elétricos, eletrônicos e de comunicação, submetidos à *Fédération des industries électriques, électroniques et de communication* (Federação da Indústrias Elétricas, Eletrônicas e de Comunicação, FIEEC).

Além do tipo anterior – ferramentas de regulação setorial –, existe a tentativa de construir cartas de ética. Essas ferramentas também deveriam ser incentivadas ao se implementar projetos dos *Big Data*. Um exemplo de atividade que poderia ser objeto de cartas de ética seriam os projetos de análise de aprendizagem, que visam explorar dados educacionais para melhorar os métodos pedagógicos no setor de educação (CHAUSSON, 2013; CAPDIGITAL, 2020). Tais cartas visam nos lembrar das obrigações impostas pela regulação relativa aos dados pessoais; mas, elas também vão além disso. Na realidade, o seu objetivo é pensar na utilização desses dados pessoais e nas suas salvaguardas. Elas poderiam resultar da análise realizada por diferentes atores da sociedade e, entre eles, uma pluralidade de partes interessadas, como advogados, analistas de sistema, e, quem sabe, filósofos ou sociólogos, a fim de fornecer uma visão completa dos usos dos *Big Data*. Essa reflexão ética parece hoje essencial para facilitar a criação, a troca e a difusão dos dados, relativamente ao respeito das liberdades fundamentais dos indivíduos. Tal reflexão deveria ser desenvolvida, e, quem sabe, generalizada, para chegar à proposta de criação de comitês de ética, semelhantes aos que têm sido criados no campo das pesquisas científicas em medicina, biologia, dentre outras áreas.

Fortalecimento dos Titulares dos Dados Pessoais

Uma vez identificados mecanismos, cabe analisar os meios para fortalecimento dos titulares dos dados pessoais no tema em questão. Os *Big Data* produzem uma forte assimetria informativa entre os responsáveis pelos tratamentos e os indivíduos. Por conseguinte, para além da responsabilização, o legislador deve igualmente assegurar a efetividade da proteção das pessoas envolvidas. Torna-se, pois, necessário promover a instauração e o reconhecimento de novos direitos para os titulares de dados pessoais e usuários de sistemas

[11] Nota do revisor técnico. O verbo iterar significa repetir. O termo é comum no jargão técnica da programação e designa repetições programadas e, portanto, codificadas.

informáticos, a fim de lhes permitir preservar os seus interesses, como será descrito na próxima subseção. Ainda, é também necessário encorajar o fortalecimento (*empowerment*) dos indivíduos, para lhes permitir exercer um contrapoder informacional, como será indicado na segunda subseção.

Novos Direitos para os Indivíduos no Contexto dos Big Data

Em primeiro lugar, devem ser concedidos novos direitos às pessoas atingidas pelo tratamento de grandes volumes de dados, para que elas possam fundamentar devidamente o seu consentimento e tomar todas as medidas necessárias para assegurar a proteção dos seus dados pessoais. Para além dos direitos já consagrados no direito de proteção de dados da União Europeia, em vigor (MATTATIA, 2016), trata-se de efetivar o direito de acesso aos elementos de perfilhamento, bem como, de se efetivar as possibilidades de retificação e de atualização de dados. Tudo isso deve ser feito para se buscar evitar ou para corrigir quaisquer erros nos tratamentos automatizados de dados pessoais. Além disso, aos titulares dos dados pessoais deve ser assegurado o direito à informação sobre as decisões que deverá tomar em função dos tratamentos de grandes volumes de dados. Ou, pelo menos, possuir o claro e efetivo direito à informação sobre a existência de um tratamento algorítmico dos seus dados pessoais, como descrito na Opinião do Grupo de Trabalho do Artigo 29 (UNIÃO EUROPEIA, 2020; UNIÃO EUROPEIA, 2013, p. 47). Para além da construção de instrumentos no cerne do Estado, há que se indicar a necessidade do fortalecimento (*empowerment*) dos indivíduos.

Big Data e a Promoção do Fortalecimento dos Indivíduos (Empowerment)

Atualmente, algumas pessoas defendem a adoção de ferramentas de capacitação destinadas a oferecer aos consumidores a oportunidade de se tornarem mais ativos no gerenciamento de seus dados pessoais. Essas soluções destinam-se a garantir a autonomia dos próprios indivíduos sobre seus dados, preconizando um modelo *user centrics*, ou seja, "centrado no usuário". Os dados pessoais do indivíduo são então centralizados em um espaço privado disponibilizado por um terceiro, prestador de serviços. Em França, esse modelo de solução tecnológica é ofertado, por exemplo, pela empresa Cozy Cloud (2020). Essas soluções de *empowerment* baseiam-se no princípio de que o indivíduo controla (vs. delega), detém (vs. centraliza) e explora (vs. compartimenta) seus dados pessoais. Uma vez capacitado para se tornar autônomo com relação à salvaguarda de seus próprios dados pessoais, o usuário estaria "habilitado" a controlar a sua utilização, podendo assim participar de sua valorização, por exemplo, graças aos sistemas do tipo *vendor relationship management* [gestão de relações com os fornecedores]. Outro movimento – derivado do movimento em

prol da expansão dos *Open Data* – é referido como *Self Data* (PUCHERAL, RALLET, ROCHELANDET e ZOLYNSKI, 2016; BÉNAVENT, 2014). Ele vai na mesma direção, ou seja, ele aponta para a produção de soluções tecnológicas que sejam capazes de ajudar na capacitação dos indivíduos em prol do seu fortalecimento como gestores dos seus dados pessoais. Algumas iniciativas emblemáticas de *Self Data* estão surgindo, tais como os projetos *Blue Button* (dados médicos) (ESTADOS UNIDOS DA AMÉRICA, 2020) e *Green Button* (dados de energia) (GREEN BUTTON ALLIANCE, 2020), nos Estados Unidos da América. No mesmo sentido estão os projetos *miData* no Reino Unido (2011) e seu equivalente *MesInfos* na França (FONDATION INTERNET NOUVELLE GÉNÉRATION, 2020). Nessas iniciativas, as pessoas são sensibilizadas para a tomada de consciência de que elas mesmas são uma fonte de dados valiosos, o que poderá encorajá-las a apoiar uma melhor proteção de seus dados pessoais. A portabilidade dos dados será uma condição para o êxito dessas propostas alternativas. O grande desafio será assegurar a interoperabilidade e a compatibilidade dos formatos de restituição e de eliminação dos dados pessoais nos servidores dos diversos prestadores de serviços (PUCHERAL *et alli*, 2016).

Qual será a eficiência dessas propostas de fortalecimento dos indivíduos (*empowerment*)? Será que eles seriam capazes de garantir uma maior proteção dos dados pessoais dos usuários? Isso pode ocorrer, desde que algumas questões sejam resolvidas antes. Há dois problemas que devem ser indicados. O primeiro se refere às questões de segurança dos dados. O segundo se refere à responsabilidade. No primeiro problema, parte-se da noção de que os dados serão centralizados em instalações sob o controle direto dos usuários e podem, portanto, ser objeto de ataques direcionados. Assim, esses ataques se tornariam, potencialmente, muito mais prejudiciais, pois o conteúdo estaria em um único local de armazenagem. O segundo problema se refere às questões de responsabilidade. Embora os indivíduos obtenham um melhor controle sobre os seus dados, também herdam a responsabilidade de protege-los. A gestão dos filtros de proteção é assim transferida dos operadores técnicos para os indivíduos que se tornaram sujeitos ativos da sua gestão de dados. Isso seria uma solução oportuna para proteger melhor os dados pessoais? Ou, ela seria uma medida para incentivar a divulgação pessoal dos dados pelos próprios usuários e, assim, estimularia o aumento do paradoxo da privacidade (*paradox privacy*) (PRAS, 2012; GERBER, GERBER e VOLKAMER, 2018)? Por fim, ainda seria necessário assegurar que essas novas soluções tecnológicas não pudessem gerar novas formas de mediação e compartilhamento de dados que, depois, pudessem se mostrar nocivas para aos próprios usuários (PUCHERAL *et alli*, 2016, p. 121). Há um longo caminho de desenvolvimento jurídico e tecnológico, como é possível compreender.

CONCLUSÃO

O presente artigo teve como objetivo central indicar os dilemas atuais que se relacionam com a expansão dos novos meios de tratamentos dos dados. O incremento na capacidade de processamento, seja por força da expansão do *hardware* ou por inovações nos *softwares*, tem gerado complexas dúvidas sobre os limites jurídicos para a oferta de soluções de proteção. O texto iniciou com a exposição dos problemas que estão sendo identificados no encaixe entre o direito vigente e as novas tecnologias de tratamento dos dados. Depois, o texto demonstrou como o Direito da União Europeia, com o advento do RGPD, trouxe princípios que são adaptáveis para a formação de um quadro jurídico protetivo. Após isso, foi necessário indicar alguns desenvolvimentos de propostas de proteção, como as soluções criadas pela CNIL, bem como soluções tecnológicas que se apresentam como possibilidades de fortalecer os indivíduos em face da grande assimetria que existe entre eles e as enormes empresas da área de Internet. É fato é que o desafio levantado pelos *Big Data* exigirá, certamente, uma reflexão coletiva – em termos de ação do Estado e da sociedade civil – para que sejam identificadas soluções que possam promover práticas de inovação responsável. Assim, o objetivo final é tentar combinar políticas regulação dos usos dos tratamentos de dados pessoais, nas melhores formas possíveis para poder se combinar os princípios de inovação e da proteção. Tal reflexão deve assumir a forma de políticas de corregulação que envolvam as mais variadas partes interessadas, bem como as autoridades reguladoras (FALQUE-PIERROTIN, GRIGUER e MOSSÉ, 2014). Há que frisar a necessidade de se criar meios hábeis para permitir a participação da sociedade civil, considerando-se, também, os grandes desafios sociais que o desenvolvimento desses meios de tratamento de dados – os *Big Data* – representam para a população como um todo, bem como, para as gerações futuras.

REFERÊNCIAS BIBLIOGRÁFICAS, LEGISLATIVAS E SÍTIOS ELETRÔNICOS

BÉNAVENT, Christophe. Big data: no best way. **Le libellio a' (AEGIS).** v. 10, n. 4, p.5-14, 2014. Disponível: http://lelibellio.com/wp-content/uploads/2013/01/Le-Libellio-d-Volume-10-num%C3%A9ro-4-Hiver-2014.pdf.

BENSAMOUN, Alexandra; ZOLYNSKI, Célia. Cloud computing et big data: quel encadrement pour ces nouveaux usages des données personnelles? **Réseaux**, n. 189, p. 103-121, 2015. Disponível: https://www.cairn.info/revue-reseaux-2015-1-page-103.htm.

BUCKLAND, Michael. *Information and society*. Cambridge, MA: The MIT Press, 2017.

CANADÁ: House of Commons. **Towards Privacy by Design: Review of the Personal Information Protection and Electronic Documents Act – Report of the Standing Committee on Access to Information**, Privacy and Ethics. Ottawa: House of Commons, 2018. Disponível: https://www.ourcommons.ca/Content/Committee/421/ETHI/Reports/RP9690701/ethirp12/ethirp12-e.pdf.

CANADÁ: Ministre de la Justice. **Loi sur la protection des renseignements personnels et les documents électroniques** *(L.C. 2000, ch. 5)*. Ottawa: Ministre de la Justice, 21 jun. 2019. Disponível: https://laws-lois.justice.gc.ca/PDF/P-8.6.pdf.

CAPDIGITAL. Charte **Ethique et Big Data:** faciiliter l'échange et la diffusion des données, 2020. Disponível: http://wiki.ethique-big-data.org/chartes/CharteEthiqueBigDataLightTCOFPOS.

CHAUSSON, Cyrille. Charte Éthique et Big Data: l'alliance Big Data veut garantir la traçabilite des données**. LeMagIT**, 7 jun. 2013. Disponível: https://www.lemagit.fr/actualites/2240200666/Charte-Ethique-et-Big-Data-lAlliance-Big-Data-veut-garantir-la-tracabilite-des-donnees.

COINTOT, Jean-Charles ; EYCHENNE, Yves**. La révolution Big data: les données au coeur de la transformation de l'enterprise.** Malakoff: Dunod Éditions, 2014.

COZY CLOUD. **Le domicile numérique pour réunir toutes vos données.** 2020. Disponível: https://cozy.io/fr.

ESTADOS UNIDOS DA AMÉRICA: Centers for Medicare & Medicaid Services. **Blue Button 2.0: A developer-friendly, standards-based API that enables Medicare beneficiaries to connect their claims data to the applications, services and research programs they trust.** Baltimore, MD: Centers for Medicare & Medicaid Services, 2020. Disponível: https://bluebutton.cms.gov.

ESTADOS UNIDOS DA AMÉRICA: The White House. **Big data:** seizing opportunities, preserving value. Washington, maio 2014. Disponível: https://obamawhitehouse.archives.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf.

FALQUE-PIERROTIN, Isabelle; GRIGUER, Merav; MOSSÉ, Marc. Comment gagner la confiance des individus à l'ère du Big data. **Cahiers de Droit de l'enterprise**, n. 6, p. 9-18, nov./dez. 2014.

FAUVARQUE-COSSON, Bénédicte (dir.); ZOLYNSKI, Célia (dir.). **Le cloud computing – l'informatique en nuage.** Paris: Société de Législation Comparé, 2014.

FONDATION INTERNET NOUVELLE GENERATION (FING). **MesInfos.** Paris: FING, 2020. Disponível: http://mesinfos.fing.org.

FRANÇA: Conseil d'Etat. **Numérique et droits fondamentaux.** Paris: La documentation française, 2014. Disponível: https://www.vie-publique.fr/sites/default/files/rapport/pdf/144000541.pdf.

FRANÇA: Commission nationale informatique et libertés. **Analyse d'impact relative à la protection des données – privacy impact assessment (PIA): la méthode.** Paris: Cnil, 2018, Disponível: https://www.cnil.fr/sites/default/files/atoms/files/cnil-pia-1-fr-methode.pdf.

FRANÇA: Commission nationale informatique et libertés. **La certification et les codes de conduite.** Paris, 19 dez. 2019. Disponível: https://www.cnil.fr/fr/la-certification-et-les-codes-de-conduite.

FRANÇA: Commission nationale informatique et libertés. La certification et les codes de conduite. Paris, 2020. Disponível: https://www.cnil.fr/fr/packs-de-conformite.

GAUTRAIS, Vincent. Proposition de règlement général sur la protection des données: un regard d'ailleurs. In: MARTIAL-BRAZ, Nathalie (dir.). **La Proposition de règlement européen relatif aux données à caractère personnel: propositions du réseau Trans Europe Experts.** Paris: Société de législation comparé, 2014. (coleção TEE, v. 9), p. 464-493.

GERBER, Nina; GERBER, Paul; VOLKAMER, Melanie. Explaining the privacy paradox: a systematic review of literature investigating privacy attitude and behavior. **Computers & Security,** v. 77, p. 226-261, ago. 2018.

GREEN BUTTON ALLIANCE. **Green Button Data**. Raleigh, NC: Green Button Alliance 2020. Disponível: http://www.greenbuttondata.org.

LATREILLE, Antoine; ZOLYNSKI, Célia. Big data et protection des données personnelles. In: MARTIAL-BRAZ, Nathalie (dir.). **La Proposition de règlement européen relatif aux données à caractère personnel: propositions du réseau Trans Europe Experts**. Paris: Société de législation comparé, 2014. (coleção TEE, v. 9), p. 262-277.

MARTIAL-BRAZ, Nathalie; ZOLYNSKI, Célia. **La gratuité: un concept aux frontières de l'économie et le droit.** Paris: LGDJ / Lextenso Éditions, 2013.

MATTATIA, Fabrice. Synthèse du futur règlement européen sur les données personnelles: principes généraux et obligations du responsable de traitement. **Révue Lamy de Droit de l'immateriel**, n. 126, p. 39-42, 2016.

MAYER-SCHÖNBERGER, Viktor; CUKIER, Kenneth. **Big data: la révolution des données est en marche**. Paris: Éditions Robert Laffont, 2014.

PRAS, Bernard. Entreprise et vie privée. *Revue française de gestion*, v. 224, p. 87-94, 2012.

PUCHERAL, Philippe; RALLET, Allain; ROCHELANDET, Fabrice; ZOLYNSKI, Célia. La privacy by design: une fausse bonne solution aux problèmes de protection des données personnelles soulevés par l'Open data et les objets connectés? **Légicom,** n. 56, p. 111-121, 2016.

RALLET, Allain; ROCHELANDET, Fabrice; ZOLYSNKI, Célia. De la privacy by design à la privacy by using. **Réseaux**, n. 189, p. 15-46, 2015. Disponível: https://www.cairn.info/revue-reseaux-2015-1-page-15.htm.

REINO UNIDO. **The midata vision of consumer empowerment.** Londres, 3 nov. 2011. Disponível: https://www.gov.uk/government/news/the-midata-vision-of-consumer-empowerment.

ROUVORY, Antoinette. Des données sans personne: le fétichisme de la donnée à caractère personnel à l'épreuve de l'idéologie des Big Data. In: FRANÇA: Conseil d'Etat. **Numérique et droits fondamentaux**. Paris: La documentation française, 2014, p. 407-421. Disponível: https://www.vie-publique.fr/sites/default/files/rapport/pdf/144000541.pdf.

SMYRNAIOS, Nikos. **Les GAFAM contre l'Internet: une économie politique du numérique**. Paris: INA, 2017.

TIC SANTÉ**. La FHF [Fédération hospitalière de France] et la Cnil planchent sur un "code de conduite" dans l'application du RGPD.** Paris, TIC Santé, 31 maio 2019. Disponível: https://www.ticsante.com/story/4629/la-fhf-et-la-cnil-planchent-sur-un-code-de-conduite-dans-l-application-du-rgpd.html. Acesso: 22 jan. 2020.

UNIÃO EUROPEIA. Regulamento 2016/679 UE, de 27 abr. 2016, define, nos Estados-membros, direitos e deveres referentes à proteção de dados pessoais e revoga a Diretiva 95/46/CE. Bruxelles: **Jornal Oficial da União Europeia**, 2016. Disponível em: http://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:32016R0679&from=PT. Acesso em: 11 out. 2017.

UNIÃO EUROPEIA: Article 29 Data Protection Working Party. **Opinion 03/2013 on purpose limitation (00569/13/EN WP 203).** Bruxelles: União Europeia, 2 abr. 2013, Disponível: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf.

UNIÃO EUROPEIA: Comité Europeu para a Proteção de Dados. **Grupo de Trabalho do Artigo 29º.** Bruxelas, 2020. Disponível: https://edpb.europa.eu/our-work-tools/article-29-working-party_pt.

WILLART, Sylvain; CRIE, Dominique. **Création de valeur par les données massives. Statistique et Société,** v. 4, n. 3, p. 19-24, dez. 2017. Disponível: https://pdfs.semanticscholar.org/36ce/dde552eb3adc950e98ea29067119bba27d60.pdf.

ZARSKY, Tal Z. Incompatible: the GDPR in the age of big data. **Seton Hall Law Review**, v. 47, n. 4. p. 995-1020, 2016.

ZOLYNSKI, Célia. **Méthode de transposition des directives communautaires: étude à partir de l'exemple du droit d'auteur et des droits voisins.** Paris: Dalloz, 2007.

ZOLYNSKI, Célia. Privacy by design appliquée aux objets connectés: vers une régulation efficiente du risque informationnel? **Dalloz IP/IT: droit de la propriété intellectuelle et du numérique**, n. 9, p. 404-480, set. 2016.

ZOLYSNKI, Célia. Big Data: pour une éthique des données. **i2D: information, données et documents,** n. 52, p. 25-26, 2015. Disponível: https://www.cairn.info/revue-i2d-information-donnees-et-documents-2015-2-page-25.htm.