REVISTA DIREITO PÚBLICO

CONTRIBUIÇÕES E LIMITES DA LEI GERAL DE PROTEÇÃO DE DADOS PARA A REGULAÇÃO DA INTELIGÊNCIA ARTIFICIAL NO BRASIL

ARTIFICIAL INTELLIGENCE REGULATION IN BRAZIL: THE CONTRIBUTION AND LIMITS OF THE GENERAL DATA PROTECTION LAW

MARCO ALMADA[1, I]
I Instituto Universitário Europeu (EUI). Florença (Toscana). Itália.

JULIANO MARANHÃO[2, II]
II Universidade de São Paulo (USP). São Paulo (SP). Brasil.

RESUMO: Sistemas de tomada de decisão automatizada são cada vez mais comuns em diversas aplicações nos setores público e privado. Como muitas das decisões tomadas por estes sistemas dependem de dados pessoais como insumos ou geram dados associáveis a pessoa natural, a Lei Geral de Proteção de Dados (LGPD) introduz regras específicas voltadas a decisões baseadas unicamente no tratamento automatizado de dados pessoais. Este artigo sustenta que o regramento das decisões automatizadas na LGPD possui contornos distintos da abordagem adotada por outras jurisdições, notadamente a União Europeia. Tais diferenças surgem das formulações adotadas pelo legislador brasileiro ao estabelecer um direito à revisão de decisões automatizadas e definir os critérios para a adoção de medidas de design para a proteção de dados, e dificultam o transplante direto de salvaguardas técnicas adotadas em outras jurisdições. A adequada proteção dos direitos do titular de dados frente às decisões automatizadas exige, portanto, o desenvolvimento de padrões e diretrizes adequados às realidades do direito brasileiro.

PALAVRAS-CHAVE: Proteção de dados; decisões automatizadas; regulação *by design*; direito à revisão; Lei Geral de Proteção de Dados (LGPD).

ABSTRACT: Automated decision-making systems are increasingly common in a broad range of public and private sector applications. Because many of the decisions made by such systems rely on personal data as input or generate data that can be associated with a natural person, the Brazilian General Data Protection Law (LGPD) introduces specific rules on decisions based solely on the automated processing of personal data. This article holds that the regulation of automated decision-making in the LGPD differs from the approaches adopted in other jurisdictions, namely the European Union. Such differences result from how Brazilian legislators established a right to the review of automated decisions and set up the criteria for the adoption of design measures for data protection, creating issues for the direct adoption of technical solutions created for other jurisdictions. Therefore, adequate protection of data subject rights against automated decision-making demands the development of standards and guidelines that reflect the realities of Brazilian law.

[1] Orcid do(a) autor(a) 1: https://orcid.org/0000-0002-0127-6549
[2] Orcid do(a) autor(a) 2: https://orcid.org/0000-0002-2705-7440

KEYWORDS: data protection; automated decision-making; data protection by design; right to review; General Data Protection Law (LGPD).

INTRODUÇÃO

Decisões automatizadas são cada vez mais comuns em vários setores da vida social. Ao navegar na internet, somos expostos a anúncios personalizados com base em perfis comportamentais (TOBBIN; CARDIN, 2020). Carros autônomos, atualmente em desenvolvimento por diversas empresas do setor automobilístico, precisarão tomar decisões a todo tempo no trânsito (ROBERTO, 2020). Sistemas de proteção ao crédito constroem *scores* e perfis mais completos de pessoas, que são então utilizados para decisões que afetarão a vida financeira destas (CITRON; PASQUALE, 2014; MARANHÃO; CAMPOS, 2019). Em todos estes casos, sistemas computacionais tomam decisões que afetam as escolhas disponíveis a seres humanos, muitas vezes sem envolvimento de qualquer ser humano no processo decisório.

Apesar de a automação, *in extremis*, eliminar o envolvimento direto de pessoas naturais em processos decisórios, a tomada de decisões automatizadas continua a ser prática com um forte componente humano. Isso porque artefatos tecnológicos não são objetos axiologicamente neutros, mas o produto de uma série de decisões, que partem de premissas culturais e políticas específicas e, assim, contribuem para consolidar certas estruturas e comportamentos sociais (WINNER, 1980). Por exemplo, sistemas de decisão automatizada são parte essencial dos modelos de trabalho da chamada economia de plataformas, direcionando os trabalhadores, precificando e supervisionando suas atividades.[3] Estes sistemas produzem, assim, efeitos que muitas vezes vão além do contexto da decisão individualizada e impactam estruturas sociais mais profundas.

A automação de processos de tomada de decisão traz consigo potenciais vantagens, como a possibilidade de reduzir o trabalho mecânico dos trabalhadores humanos e de garantir uma prestação mais eficiente de produtos e serviços. Tais vantagens, contudo, são acompanhados por uma série de riscos para indivíduos e para a sociedade. Em particular, há o risco de que decisões automatizadas baseadas em perfis, construídos a partir de dados pessoais e analisados por complexos algoritmos, produzam resultados discriminatórios ou não suficientemente transparentes para aqueles afetados pela decisão. Objetivando especificamente

[3] Ver, entre outros, (TOMASSETTI, 2020)

proteger titulares de dados em relação ao processamento de seus dados para decisões automatizadas que os afetem, a Lei Geral de Proteção de Dados (LGPD)[4] traz o Artigo 20, que provê às pessoas naturais alguns direitos oponíveis ao uso de seus dados pessoais para a tomada de decisões automatizadas: o direito de revisão, explicação e a possibilidade de auditoria pela autoridade para verificação de potencial discriminatório.

Contudo, as tecnologias hoje utilizadas para automação, em especial aquelas baseadas em técnicas de aprendizado de máquina, incorporam diferentes tipos e níveis de opacidade (BURRELL, 2016), que geram obstáculos ao exercício destes direitos: como revisar uma decisão baseada em uma função probabilística complexa pela qual não fica claro quais fatores determinaram seu resultado? E ainda quando for possível encontrar correlações estatisticamente significantes entre variáveis de entrada e a predição de saída, como traduzi-las em relações causais ou finalísticas compreensíveis por humanos? Como pleitear a revisão quando sequer se sabe da existência de uma decisão automatizada? Para evitar que a complexidade técnica se torne um escudo para evitar a aplicação da lei,[5] a LGPD inclui em seu artigo 46 a chamada proteção de dados *by design*, segundo o qual os agentes de tratamento de dados pessoais têm o dever de adotar medidas técnicas, administrativas e de segurança para assegurar que dados pessoais não sejam tratados de forma inadequada ou ilícita. Porém, dada a vasta gama de aplicações possíveis para a tomada automatizada de decisão, cada uma com seus riscos associados, tais dispositivos adotam uma redação aberta, que traz poucos detalhes a respeito do conteúdo específico dos deveres de *design* e de revisão.

Como tal abertura normativa pode trazer dificuldades para a aplicação dos dispositivos em questão, o presente artigo propõe uma interpretação sistemática destes como instrumentos para a regulação de risco.[6] Uma vez que a LGPD estabelece a prevenção de danos como um de seus princípios-guia (LGPD, art. 6º, VIII), os agentes de tratamento de dados devem avaliar os potenciais danos que posam decorrer de cada aplicação de decisão automatizada e adotar medidas que minimizem os riscos de concretização destes danos, bem como tornem viável a

[4] (BRASIL, 2018)
[5] Ver, por exemplo, (BRYSON; DIAMANTIS; GRANT, 2017)
[6] É importante frisar que os riscos não fornecem a única baliza da LGPD, que também estipula a defesa de diversos direitos individuais e coletivos ligados ao tratamento de dados pessoais (ZANATTA, 2021, seç. 2.1). Tais direitos restringem a margem de discrição do agente regulado, mas os deveres deste ainda são formulados em termos de riscos: ver, e.g., o Artigo 44 da LGPD. Para um tratamento mais geral do papel da regulação de risco na proteção de dados, ver (GELLERT, 2018).

revisão de eventuais decisões que produzam efeitos prejudiciais aos interesses do titular de dados. O diagnóstico dos riscos é, por definição, dependente do contexto em que a decisão automatizada ocorre. Contudo, é possível extrair do texto da LGPD algumas balizas para esta avaliação contextual, o que é feito no restante deste artigo.

Na próxima seção, o artigo analisa os contornos normativos do direito à revisão de decisões automatizadas como posto no Artigo 20 da LGPD, contrastando-o com seu análogo no ordenamento da União Europeia (Artigo 22 do RGPD[7]). Aqui, o ordenamento brasileiro inova ao estabelecer um direito de revisão e não uma restrição às possibilidades de uso de decisões automatizadas, mas segue a tendência internacional ao estender tal direito apenas às decisões tomadas sem envolvimento humano. A seção seguinte examina as medidas de proteção de dados *by design* incorporadas ao ordenamento jurídico brasileiro pelos Artigos 46 e 49 da LGPD, sustentando que elas podem não só contribuir para a efetividade do direito à revisão como oferecer salvaguardas nos casos de decisões apenas parcialmente automatizadas. Diante da indeterminação associada ao controle de riscos decorrentes de tecnologias que podem ser usadas para diversos propósitos, a seção seguinte examina como a Autoridade Nacional de Proteção de Dados (ANPD) pode contribuir para a efetiva aplicação do direito à revisão e da proteção de dados *by design*. Com isso, pretende-se mostrar tanto a existência de abordagens técnicas para a governança das tecnologias de automação de decisão quanto a possibilidade de que as ferramentas jurídicas para tal governança sejam construídas a partir da cooperação entre a atuação regulatória estatal e a inovação dos atores públicos e privados na construção de sistemas que se valham dos avanços tecnológicos ao mesmo tempo que asseguram a proteção dos dados pessoais.

AS DECISÕES AUTOMATIZADAS NA LGPD

A ideia de que deve existir alguma possibilidade de controle humano sobre decisões que, em regra, são tomadas de forma automatizada encontra eco em várias das legislações de proteção de dados, como o RGPD e a LGPD. No caso específico do Brasil, a formulação vigente do *caput* do Artigo 20 da LGPD estipula que o titular de dados tem direito a

> [...] solicitar a revisão de decisões tomadas *unicamente* com base em tratamento automatizado de dados pessoais que afetem seus interesses, incluídas as decisões

[7] Regulamento Geral de Proteção de Dados (RGPD), também conhecido pela sigla em inglês, GDPR (UNIÃO EUROPEIA, 2016).

destinadas a definir o seu perfil pessoal, profissional, de consumo e de crédito ou os aspectos de sua personalidade.

O direito à revisão permite, portanto, a alteração do desfecho de um processo decisório que seja inteiramente conduzido por máquina. Quando esta decisão é inteiramente informacional, ou seja, tem como resultado uma decisão que não produz impactos diretos em um ambiente físico—por exemplo, uma decisão que determina se uma pessoa tem ou não direito a receber um benefício social—este direito se assemelha ao direito à correção dos dados pessoais estabelecido pela LGPD no Artigo 18, III, pois em ambos os casos, o titular tem o direito solicitar ao agente de tratamento de seus dados pessoais a alteração de uma informação que diz respeito a ele e que, a seu ver, é errônea. Diante da existência de tal direito, a revisão dos fatos associados a uma decisão automatizada, ou produzidos por uma, independe de um direito à revisão.

Não se pode concluir, todavia, que o direito à revisão careça de conteúdo próprio. Em muitos casos, a automação não visa apenas a produção de um perfil comportamental ou outra forma de extração de informação ao respeito do titular de dados, mas sim a produção de um desfecho no mundo real, sem a participação direta de um humano. Por exemplo, um sistema de triagem de currículos processa dados pessoais, e para isso realiza uma avaliação do titular desses dados. A decisão propriamente dita, nesse caso, não é a predição de comportamento do indivíduo, mas uma ação—permitir ou não que o candidato siga no processo seletivo—a ser valorada à luz do ordenamento jurídico. Uma ação, ao contrário de uma informação, não é algo que possa ser classificado como verdadeiro ou falso (BEIGANG, 2022), mas como correta ou incorreta, boa ou ruim, lícita ou ilícita, adequada ou inadequada para determinado fim. Portanto, o direito à correção teria pouco a oferecer ao titular de dados quanto aos efeitos da decisão, ao passo que a revisão traz consigo a possibilidade de reversão do curso de ação original.

Ou seja, embora a correção da informação possa afetar o processo decisório, o direito à retificação tem por objeto assegurar a acurácia do dado armazenado, ou seja, do *produto* de operações computacionais. Já no direito à revisão, o dever correlato tem por objeto o *processo decisório* pelo qual a decisão automatizada é tomada. Enfatizamos o processo, pois o direito à revisão não só pauta o desfecho ou resultado—afinal, um revisor humano pode chegar à mesma conclusão—mas pretende assegurar a adequação de todas as etapas que levaram a uma

determinada decisão e que deve ser entendida com base em seus propósitos e elementos constitutivos elencados no texto legal.[8]

A terminologia adotada no *caput* do artigo 20 da LGPD ressalta, de pronto, alguns dos contornos postos pelo legislador à revisão das decisões automatizadas. O estabelecimento de um direito do titular apenas à *revisão* de decisões já é, em si, uma escolha significativa, pois reflete uma autorização implícita de qualquer tratamento automatizado de dados pessoais que satisfaça as exigências da lei, dentre elas a necessidade de ocorrência de uma das hipóteses de tratamento do artigo 7º da LGPD, ou, no caso de dados sensíveis, ao art. 11 da LGPD. Se o titular tem direito a solicitar revisão de decisões que lhe afetem e sejam baseadas unicamente no tratamento automatizado de dados pessoais, então já se pressupõe uma permissão implícita ao controlador para processamento de decisões automatizadas baseadas em dados pessoais. Desse modo, a permissão, mesmo que implícita, é mais forte do que uma simples ausência de proibição, refletindo um direito do controlador.

Ainda que parte da doutrina europeia interprete o Artigo 22 do RGPD em linhas similares à brasileira, a prática judicial e administrativa na União Europeia tende a adotar interpretação mais restritiva, em linha com a literalidade do § 1° deste artigo, que se refere a um direito a não ser submetido a decisões automatizadas (VALE; ZANFIR-FORTUNA, 2022). Esta interpretação restritiva, vista em termos das posições jurídicas hohfeldianas (HOHFELD, 1913), é diametralmente oposta à abordagem brasileira, pois constrói o Artigo 22 como estabelecendo uma proibição da adoção de decisões automatizadas. O modelo brasileiro de proteção de dados pessoais se mostra, assim, mais permissivo em relação ao uso da automação do que seus análogos estrangeiros, uma vez que adota apenas controles *post hoc* sobre decisões automatizadas, por meio dos direitos do titular de revisão, explicação e auditoria sobre discriminação.

Outra diferença importante entre a legislação de proteção de dados brasileira e aquela adotada na União Europeia está no tipo de decisões cobertas pelo direito à revisão. Na LGPD (artigo 20, *caput)*, o direito de revisão pode ser exercido contra tratamento automatizado de dados pessoais que afete os *interesses* do titular de dados. Tal formulação cobre uma ampla gama de interesses juridicamente protegidos (LÓPEZ, 2020; MARANHÃO; ALMADA, 2021).

Na União Europeia, o direito a não ser submetido a decisões automatizadas tem uma hipótese mais estreita: tal direito se aplica apenas às decisões que produzam efeitos jurídicos ou cujos efeitos sejam significativos de forma similar a uma decisão jurídica, por exemplo ao resultar na perda de um emprego (ARTICLE 29 WP, 2018; RECHTBANK AMSTERDAM, 2021). Ainda que exista uma tendência nos tribunais e autoridades administrativas a interpretar o critério de significância do RGPD de forma ampla (VALE; ZANFIR-FORTUNA, 2022, p. 35–38), a adoção de um critério de relevância significa que nem todas as decisões automatizadas são cobertas pelo Artigo 22 do RGPD. Já na LGPD, qualquer interesse juridicamente reconhecido pode dar margem a um pedido de revisão, ao menos neste primeiro momento em que não há restrições vindas da jurisprudência ou de instrumentos regulatórios. A legislação brasileira impõe, portanto, um escopo mais amplo para a aplicação do direito à revisão das decisões automatizadas.

A ampliação do escopo do direito à revisão foi acompanhada por uma flexibilização do processo revisional: enquanto o RGPD determina que o titular de dados pode pleitear a intervenção de uma pessoa natural no processo decisório, a LGPD não especifica que a revisão deva ser feita por humano. Tal especificação estava presente na redação da lei originalmente aprovada pelo Congresso, mas foi suprimida em meio às alterações realizadas pela Medida Provisória 869/2018. Ao votar a conversão desta MP em Lei, o Congresso Nacional incluiu no texto do artigo 20 o § 3°, segundo o qual a revisão deveria ser efetuada por pessoa natural, em termos a serem disciplinados pela ANPD. Esta adição, contudo, foi objeto de veto presidencial, fundamentado no potencial de que a obrigação de revisão humana torne inviável diversos modelos de negócio de empresas que usam sistemas de decisão automatizada, como instituições financeiras (PRESIDÊNCIA DA REPÚBLICA, 2019). Ainda que, na apreciação dos vetos à lei de conversão, ambas as casas do Congresso tenham votado em favor da restauração do requisito, tal voto não atingiu o requisito de maioria absoluta para a derrubada do veto (SILVA, P.; MEDEIROS, 2019). Conclui-se, portanto, que o legislador acabou por rejeitar a definição do direito à revisão como um direito à revisão humana.

Por fim, outra dúvida interpretativa trazida pela redação do Artigo 20 da LGPD está no *status* do perfilamento (*profiling*). Tanto a LGPD quanto o RGPD fazem referência ao perfilamento como uma prática associada às decisões automatizadas, mas a legislação brasileira não define o conceito de "perfil", ainda que a expressão "perfil comportamental" apareça no

Artigo 12 da LGPD. A falta de uma definição explícita pode trazer dificuldades jurídicas no futuro, ainda que a experiência de outros campos do direito—como a regulação do crédito—forneça diversos exemplos de práticas abarcadas pelo perfilamento. É importante salientar, todavia, que o Artigo 20 da LGPD estende o direito de revisão também às "decisões tomadas unicamente com base em tratamento automatizado de dados pessoais, *incluídas as decisões destinadas a definir o seu perfil pessoal* [itálico nosso]". Ou seja, o direito à revisão seria imponível não só às decisões que usam perfis como aporte decisório, mas também àquelas envolvidas na formação do perfil propriamente dito. Por exemplo, uma decisão humana sobre concessão de crédito baseada em *scoring* individual poderia ter seu score revisado, a partir de explicações trazidas pelo controlador ao titular sobre o modelo de classificação.

Soluções adotadas por diferentes ordenamentos jurídicos refletem as diferenças entre os contextos políticos, legais e sociais em que cada legislador opera. Assim, a mera adoção acrítica de uma abordagem referendada em outro ordenamento pode ser insuficiente para contemplar as necessidades postas pela legislação brasileira. Por exemplo, um protocolo de controle humano de decisões automatizadas baseado no modelo europeu adotará controles *ex ante*—como a proibição de decisões automatizadas por padrão—que não são exigidos pela lei brasileira. Ao mesmo tempo, tal protocolo enfrentará dificuldades para lidar com o escopo mais amplo e potencialmente mais volumoso de revisão permitido pelo Artigo 20 da LGPD, em relação ao do RGPD. Neste contexto, o mapeamento das diferenças entre o direito à revisão e as abordagens estrangeiras é uma tarefa importante, para a qual este artigo presta uma contribuição inicial.

Não é possível concluir, contudo, que a abordagem brasileira para decisões automatizadas seja inteiramente distinta do exemplar estrangeiro. Uma primeira semelhança relevante para os fins deste artigo está ligada ao *propósito* do dispositivo: tecnologias de automação são usadas para uma variedade de aplicações, às quais o direito à revisão e seus análogos buscam dar uma resposta jurídica. Examinaremos a seguir as justificativas levantadas na literatura internacional para a revisão de decisões automatizadas, encontrando suas bases jurídicas na LGPD.

A primeira justificativa para o direito à revisão e seus análogos está ligada à possibilidade de erros no processo de tomada de decisão automatizada. O uso do termo *revisão* pode abranger diferentes motivações para a inconformidade com o resultado da decisão: o

sujeito do dado pode pedir a revisão em função de *erro* ou por mesmo por discordar de suas implicações.

No caso de erro, há duas hipóteses a serem consideradas. A primeira delas refere-se a situações em que dados errôneos ou de baixa qualidade são utilizados para a construção do sistema de tomada de decisão ou para a tomada de uma decisão em particular. Neste caso, o titular dos dados encontra proteção imediata no Artigo 6°, V, da LGPD. Contudo, nem todos os erros de um sistema são produzidos por dados errôneos: mesmo a mais fidedigna base de dados pode não impedir falhas na modelagem técnica do problema ou na execução do *software* responsável pelas decisões.[9] Uma vez que o tratamento de dados pessoais é, também, pautado pelo princípio da prevenção (LGPD, Artigo 6º, VIII), cabe aos agentes de tratamento assegurar que fatores deste segundo tipo não resultem em danos aos titulares de dados.[10]

O caso de discordância do conteúdo da decisão, independentemente de erro material específico, por sua vez, precisa ser avaliado, considerando-se que o objeto da decisão pode envolver aspectos valorativos. Ainda que, à primeira vista, um sistema computacional possa parecer uma representação neutra, o processo de desenvolvimento desse sistema invariavelmente envolve decisões a respeito de como interpretar determinados valores e princípios (ALMADA, 2023). Os efeitos valorativos de decisões técnicas são objeto de ampla discussão na literatura a respeito da inteligência artificial, mas quaisquer conclusões são contextuais e passíveis de mudança com o passar do tempo (GORDON; RIEDER; SILENO, 2023). Tal caráter contingente dos valores materializados em sistemas de IA, bem como a própria natureza política de tais direitos faz com que a possibilidade de contestação dos resultados algorítmicos seja defendida mesmo em casos em que uma decisão automatizada possa apresentar desempenho superior à de um decisor humano (HILDEBRANDT, 2019).

O direito a contestar decisões automatizadas é, em parte, assegurada pela possibilidade de tutela em juízo dos interesses e direitos dos titulares de dados pessoais, mas Claudio Sarra (2020, p. 8) propõe que a contestação deve ser entendida como um processo dialógico mais amplo, que ofereça ao titular de dados a oportunidade de questionar formalmente eventuais

[9] A respeito dos diversos modos de falha de sistemas automatizados, e dos perigos de pressupor que a operação destes sempre leva a resultados corretos, ver (RAJI *et al.*, 2022).

[10] Estes danos podem decorrer, por exemplo, da produção de decisões inadequadas em função de falhas no algoritmo, ou de vazamento de informações a respeito do titular de dados por falhas de segurança no processo decisório.

violações a seus direitos e interesses. O mecanismo de revisão oferecido pelo Artigo 20 da LGPD desempenha este papel, ao habilitar o titular de dados a questionar seja o desfecho da decisão, seja o processo decisório em si.[11] Desta forma, o exercício do direito à revisão se mostra como uma via extrajudicial[12] para a contestação de decisões automatizadas, mesmo que o termo "contestação" não seja diretamente invocado pelo legislador brasileiro neste contexto. Como indicado acima, a efetividade a abrangência da contestabilidade de decisões automatizadas liga-se diretamente à abrangência, profundidade e qualidade da explicação sobre a decisão, critérios e processo decisório a que tenha acesso.

O direito à revisão surge, portanto, como uma forma de articular e promover estes três interesses—a *qualidade dos dados* e dos *outputs* automatizados, a proteção de *valores* afetados por processos de automação, e a possibilidade de *contestação* dos resultados de uma decisão automatizada sem que seja necessário recorrer a vias judiciais—, que devem ser avaliados no caso concreto da decisão cuja revisão se pleiteia. A revisão das decisões automatizadas surge como um instrumento para a promoção da qualidade dos processos decisórios, para a efetivação do direito à contestação destes e para a proteção da dignidade das pessoas afetadas pelas decisões. Sua disciplina e implementação, portanto, deve ter estes *desiderata* como balizas.

O alcance deste direito, assim como de seus correspondentes em outros ordenamentos, é limitado por sua forma jurídica. A LGPD regula decisões automatizadas ao estabelecer que o titular dos dados pessoais tem um *direito* a solicitar revisão, que é fraseado em termos da proteção dos dados pessoais do indivíduo diante de decisões automatizadas. Tal enquadramento traz dois limites fundamentais a este instrumento. O primeiro é que ele não se aplica a decisões que sejam apenas parcialmente automatizadas, isto é, nas quais um humano tome a decisão última com base em aportes providos por meios automatizados ou sistemas que não tenham como *output* uma decisão[13]. Nestes casos, a proteção dos titulares de dados face às decisões automatizadas é garantida pela aplicação de outros dispositivos da LGPD, bem como por uma

[11] Ainda que, na prática, seja implausível esperar um elevado número de pedidos de revisões baseados puramente em fatores procedimentais e não em um desfecho indesejável para o titular de dados.
[12] Já que a possibilidade de defesa de interesses individuais e coletivos em juízo é protegida pelo art. 22 da LGPD.
[13] Por exemplo, a definição de decisões automatizadas não engloba as decisões produzidas por sistemas de recomendação, como aqueles que indicam notícias em redes sociais, ainda que estes possam afetar os interesses do titular de dados de formas similares às decisões inteiramente automatizadas (COBBE; SINGH, 2019). Também pode ser difícil delimitar na prática quando uma decisão conta com envolvimento humano (BINNS; VEALE, 2021).

interpretação extensiva dos critérios para que uma decisão seja considerada automatizada.[14] O segundo diz respeito à aparente dimensão individual do direito à revisão ao restringir o agente ao "titular", o que torna a proteção insuficiente para capturar as dimensões transindividuais dos danos decorrentes de decisões automatizadas, como, por exemplo, aquelas que transbordam a individualidade em discriminações algorítmicas (ver, por exemplo, ABREU, 2018; VAN BEKKUM; BORGESIUS, 2023). Desta forma, o direito à revisão não cobre todas as implicações jurídicas das decisões automatizadas, mas apenas aquelas que se traduzam em dano a interesses individuais.

INCORPORANDO A PROTEÇÃO DE DADOS *BY DESIGN* ÀS DECISÕES AUTOMATIZADAS

A crescente automação de processos de tomada de decisão traz consigo questões sobre como assegurar a legalidade e a segurança das decisões automatizadas. Ao longo das últimas décadas, cientistas da computação e outros profissionais trabalharam no desenvolvimento das chamadas tecnologias de melhoria da privacidade, isto é, tecnologias voltadas a proteger e até mesmo fortalecer a privacidade de seus usuários.[15] A ideia central por trás destas tecnologias—aproveitar as capacidades tecnológicas para proteger direitos—logo encontrou respaldo em meio à pesquisa jurídica e às práticas legislativas, que passaram a recomendar e até mesmo exigir a adoção de medidas voltadas a assegurar a proteção dos dados pessoais (BYGRAVE, 2017; TRIBUNAL EUROPEU DE DIREITOS HUMANOS, 2008). No Brasil não foi diferente, uma vez que o texto da LGPD acolheu a chamada proteção de dados *by design*, obrigando assim os agentes de tratamento de dados pessoais a usar a tecnologia para cumprir seus deveres em relação aos titulares de dados (RIVELLI; SILVEIRA, 2021).

O núcleo da proteção de dados *by design* no ordenamento brasileiro está no Artigo 46 da LGPD. Em seu *caput*, este artigo impõe que os agentes de tratamento de dados pessoais devem adotar medidas de segurança, técnicas e administrativas para a proteção de dados pessoais. Ao contrário do que ocorre no dispositivo europeu análogo (RGPD, Artigo 25(1)), a LGPD não definiu tais medidas em função dos riscos ligados ao tratamento de dados pessoais, exigindo em vez disso que as medidas adotadas sejam aptas a proteger dados pessoais contra

[14] A respeito da aplicação dos dispositivos de decisões automatizadas por autoridades judiciais e administrativas na União Europeia, ver (VALE; ZANFIR-FORTUNA, 2022).
[15] Estas tecnologias também são conhecidas como PETs, a partir da sigla em inglês *privacy-enhancing Technologies* (TAMÒ-LARRIEUX, 2018, pp. 21-22).

os fatores elencados no dispositivo.[16] Todavia, a necessidade de adotar medidas que previnam riscos decorrentes da proteção de dados segue do fato de que a prevenção é um princípio geral do regime brasileiro de proteção de dados (LGPD, Artigo 6, VIII). A aptidão referida no *caput* do Artigo 46 da LGPD deve, portanto, ser avaliada em termos da eficácia de um conjunto de medidas em prevenir—ou ao menos mitigar—os riscos decorrentes do tratamento de dados pessoais (MARRAFON; COUTINHO, 2020).

Decisões automatizadas, quando baseadas no tratamento de dados pessoais, introduzem riscos específicos a serem enfrentados nos termos do *caput* do Artigo 46 da LGPD. Estes riscos, que se somam àqueles envolvidos à proteção de dados de uma forma geral, estão ligados aos mesmos interesses que guiam o direito à revisão, descritos acima. Uma decisão enviesada ou que seja de outra forma de má qualidade pode impactar os direitos do titular de dados, pois será baseada em dados que não representam adequadamente a realidade ou as populações afetadas. Uma decisão que não proteja a dignidade do titular de dados pode afetar não só seu direito à proteção de dados pessoais, mas outros direitos fundamentais que sejam impactados pelos efeitos da decisão.[17] Muitos destes impactos vão além da esfera da autodeterminação informativa que sustenta o direito à proteção de dados (MENDES, 2020), produzindo impactos em outros interesses do indivíduo ou mesmo em outros de seus direitos fundamentais. Não à toa, o *caput* do artigo 46 da LGPD exige que as medidas técnicas, administrativas e de segurança sejam aptas a prevenir qualquer forma de tratamento inadequado ou ilícito, não só aquelas que violem o disposto naquela lei.[18]

Tais medidas de proteção do titular de dados podem tomar diversas formas. Embora o conteúdo destas medidas não seja explicitamente definido no texto da LGPD, é possível delimitá-lo a partir do significado dos termos e dos aportes das experiências internacionais. Medidas técnicas consistem na adoção de tecnologias que reduzam ou eliminem riscos ligados ao tratamento, como a anonimização dos dados a serem tratados.[19] Já as medidas

[16] Nos termos do *caput* do Artigo 46 da LGPD, estes são "acessos não autorizados e de situações acidentais ou ilícitas de destruição, perda, alteração, comunicação ou qualquer forma de tratamento inadequado ou ilícito".
[17] Uma decisão que não abra margem para contestação viola, no mínimo, o direito que o titular de dados pessoais tem de recorrer à justiça para reivindicar seus direitos.
[18] Tal leitura segue da definição do princípio da prevenção (LGPD, Artigo 6º, VIII), segundo o qual se deve prevenir que o tratamento de dados resulte em danos, não restringindo estes aos danos ligados à autodeterminação informativa.
[19] Para uma introdução à anonimização, os riscos aos quais ela responde, e seus limites, ver (FINCK; PALLAS, 2020).

administrativas, também conhecidas como organizacionais,[20] referem-se ao contexto institucional em que um sistema é utilizado.[21] Por fim, as medidas de segurança englobam tanto medidas técnicas quanto administrativas, sendo destacadas pelo legislador em função de seu objeto: a prevenção de acessos não autorizados e da destruição, perda, alteração, comunicação ou difusão dos dados (LGPD, Artigo 6º, VII).

Na prática, uma abordagem efetiva para a proteção de dados pessoais combinará abordagens técnicas e administrativas, já que alguns problemas são mais bem resolvidos por um meio do que por outro. Por exemplo, as autoridades de proteção de dados da União Europeia recentemente recomendaram que a transferência de dados pessoais para os Estados Unidos seja acompanhada pelo uso de medidas técnicas fortes, em particular a anonimização dos dados enviados.[22] No sentido oposto, existem situações em que uma medida administrativa pode ser mais eficaz que soluções técnicas na resolução de um dado problema. Por exemplo, uma empresa de pequeno porte, a depender de suas circunstâncias, pode prescindir de controles técnicos sofisticados caso adote políticas de acesso adequadas e assegure que os profissionais que lidarão com os dados sejam adequadamente treinados. Dados os limitados riscos aos direitos dos titulares de dados e escala da operação, adotar o tipo de controle necessário em uma grande empresa seria muito oneroso e não traria proteção adicional. A definição das medidas a serem adotadas depende, portanto, do cenário em que ocorre o tratamento de dados, que determinará os riscos decorrentes da aplicação das tecnologias disponíveis para tal tratamento.

Uma vez que o contexto de operação de sistemas de decisão automatizada—de fato, de qualquer sistema de tratamento de dados pessoais—muda com o tempo, a adequada proteção dos titulares de dados pode requerer ajustes às práticas estabelecidas. Ciente deste fato, a LGPD estabelece que a proteção de dados *by design* é um dever contínuo: os agentes de tratamento de dados pessoais devem adotar medidas técnicas, administrativas e de segurança desde a fase da concepção do produto ou serviço até a sua execução (LGPD, art. 46, § 2°). Ou seja, os agentes

[20] É o termo usado na União Europeia (RGPD, Artigo 25(1)), mas também aparece na LGPD: Artigo 45, § 5º.

[21] Algumas destas medidas têm natureza educacional, como a adoção de treinamentos para os funcionários que utilizarão um sistema corporativo. Outras se referem aos arranjos internos para o tratamento de dados, como a definição de competências para o uso do sistema. Um terceiro conjunto de medidas administrativas trata de estruturas que facilitem o exercício dos direitos do titular de dados, como a criação de canais dedicados de atendimento.

[22] Dados os amplos poderes que a legislação americana confere às autoridades em termos de requisição de dados para fins de segurança nacional, as autoridades europeias julgaram que nenhuma medida organizacional bastaria para proteger os dados contra acessos que não seriam aceitáveis nos parâmetros da lei europeia (EDPB, 2020).

de tratamento de dados pessoais são obrigados a levar em conta as características das técnicas que utilizarão para tratar dados, implementar quaisquer medidas que enfrentem os riscos decorrentes destas técnicas, e atualizar tais medidas conforme elas percam eficácia ou se tornem obsoletas.

No caso de um sistema de decisão automatizada, isso significa que o direito à revisão—que, nos termos do *caput* do Artigo 20, é essencial à licitude de uma decisão automatizada—não pode ser visto como um enxerto posterior ao processo decisório. Em vez disso, a revisão e outros direitos do titular de dados devem ser levados em conta já a partir da escolha dos meios que serão usados para automatizar uma dada solução.[23] Ademais, estes agentes de tratamento devem adotar medidas que permitam detectar problemas com os processos decisórios, como a ocorrência de padrões discriminatórios nas decisões (WACHTER; MITTELSTADT; RUSSELL, 2021) e, uma vez detectados tais problemas, adotar medidas técnicas e administrativas para reduzir ou evitar seu impacto. Surge, portanto, a questão de como identificar e implementar tais medidas em sistemas existentes.

Os requisitos de *design* postos na LGPD não se restringem, contudo, à implantação do direito à revisão de decisões automatizadas. Uma vez que estes dispositivos exigem não só o cumprimento da legislação de proteção de dados, mas a prevenção de danos decorrentes de tratamento ilícito ou inadequado dos dados pessoais, os agentes de tratamento de dados pessoais são obrigados a adotar medidas também nos casos das decisões parcialmente automatizadas descritas acima. Mesmo que tais decisões não preencham os critérios para a aplicabilidade do direito à revisão, elas ainda usam dados pessoais em formas que podem causar danos aos usuários.[24] Como consequência, os agentes de tratamento de dados pessoais devem identificar os riscos associados à automação parcial e adotar medidas aptas a preveni-los, ainda que estas medidas não exijam a revisão da decisão, posto que esta pode ser feita pelo humano envolvido no processo decisório.

Uma categoria particularmente relevante de risco em sistemas de IA é aquela dos riscos decorrentes da *opacidade* de um sistema. Dada a complexidade envolvida nas técnicas de inteligência artificial, que muitas vezes fazem uso de sofisticados procedimentos matemáticos

[23] O que pode levar, por exemplo, à adoção de tecnologias de inteligência artificial explicável como uma forma de garantir o acesso à informação previsto pelo art. 20, § 1°, da LGPD (MARANHÃO; COZMAN; ALMADA, 2021).
[24] Ver, entre outros, (ABREU, 2018; COBBE; SINGH, 2019; SILVA, T., 2020).

aplicados a grandes volumes de dados (LAGIOIA; SARTOR, 2020), muitas vezes é inviável que um indivíduo—por mais tecnicamente capacitado que seja—entenda o que se passa de um algoritmo. Por exemplo, pode ser difícil identificar qual fator levou à rejeição de um pedido de crédito se essa decisão levou em conta centenas de variáveis,[25] selecionadas por milhões de observações de correlações em um processamento computacional.[26] Na ausência de tais informações, indivíduos afetados por decisões automatizadas podem ser privados da possibilidade de contestar uma decisão automatizada ou mesmo de saber da existência de uma decisão.[27]

Neste contexto, o uso de técnicas de inteligência artificial explicável aparece como uma solução para, ao menos, mitigar os riscos decorrentes da opacidade (MULHOLLAND; FRAJHOF, 2019). A pesquisa neste campo é fenômeno relativamente recente, mas produziu diversas abordagens que prometem reduzir a complexidade das tecnologias de inteligência artificial a um nível que possa ser entendido por indivíduos,[28] inclusive com a divulgação de padrões técnicos auditáveis para introdução de métodos de explicação (WINFIELD *et al.*, 2021). Mesmo que alguns estudos recentes apontem limites às capacidades de gerar explicações que possam substituir a análise direta de um sistema algorítmico (BORDT *et al.*, 2022), os resultados atingidos pelos métodos existentes sugerem o potencial de abordagens técnicas em criar condições que tornem um sistema de IA mais legível a seus usuários e às pessoas afetadas por sua operação (MALGIERI; COMANDÉ, 2017). A adoção de técnicas de explicação e similares pode, portanto, facilitar o exercício de direitos como o direito à revisão, ao reduzir as barreiras informacionais diante dos titulares de dados.

---

[25] Aqui, tratamos apenas das variáveis de fato empregadas em um processo decisório. Este valor não deve ser confundido com o número muito maior de variáveis usado para a construção de modelos de grande escala, como o ChatGPT, uma vez que estas se manifestam no estágio anterior de criação do modelo, suscitando problemas de responsabilização que são distintos daqueles diretamente presentes na decisão automatizada: (COBBE; VEALE; SINGH, 2023).

[26] Ver, por exemplo, o uso de dados de telefonia para a geração de *scores* de crédito (DIETTRICH; DE SOUZA; GUERREIRO, 2020).

[27] Os obstáculos criados pela opacidade técnica não são absolutos, uma vez que técnicas investigativas podem revelar eventuais problemas com algoritmos mesmo quando os parâmetros e o uso destes são mantidos opacos (WASHINGTON, 2016). Todavia, tais medidas demandam um substancial esforço técnico para a reconstrução dos fatores desconhecidos, e sua efetividade pode ser limitada por fatores como sigilo comercial e industrial (BUSUIOC; CURTIN; ALMADA, 2023). Ainda que o Artigo 20, § 2º, da LGPD confira à ANPD o poder para realizar auditorias quando tais formas de sigilo são invocadas, eventuais auditorias são limitadas à verificação de aspectos discriminatórios, e precisarão elas mesmas examinar os fatores técnicos em questão.

[28] A respeito das técnicas de inteligência artificial explicável e sua conexão com requisitos jurídicos, ver (BIBAL *et al.*, 2021; MARANHÃO; COZMAN; ALMADA, 2021).

A proteção de dados *by design* também trata dos danos sistêmicos decorrentes de decisões total ou parcialmente automatizadas. Como estes efeitos não decorrem de um tratamento em específico, mas sim do agregado de múltiplas operações cobertas (SCHAFER, 2021), eles não são cobertos pelo *caput* do artigo 46, cuja exigência é voltada às operações individuais de tratamento. Todavia, o artigo 49 da LGPD estabelece que os sistemas usados para o tratamento de dados "devem ser estruturados de forma a atender aos requisitos de segurança, aos padrões de boas práticas e de governança e aos princípios gerais previstos nesta Lei e às demais normas regulamentares." Dada a construção ampla do princípio da prevenção na LGPD, o cumprimento deste dever exige que a estrutura dos sistemas de tratamento de dados leve em conta os danos causados pela operação do sistema como um todo, não só nas operações de tratamento tomadas individualmente.

A escolha de medidas técnicas e administrativas para sistemas de decisões automatizadas pode extrair lições da experiência com a chamada *privacy by design* (PbD), termo usado para designar um conjunto de técnicas no qual a privacidade dos titulares de dados é tratada como um objetivo central do processo de construção de *software* (LIMA; ALMADA; MARANHÃO, 2022, seç. 2.2). Ao tratar a privacidade como um requisito para que um projeto de sistema seja considerado bem-sucedido, a abordagem PbD força os desenvolvedores e usuários de um sistema computacional a considerar como as demandas decorrentes da proteção da privacidade devem ser tratadas em cada estágio do ciclo de vida do sistema: desde a sua concepção até o fim de sua operação, o que inclui a programação do sistema, seus testes e modos de uso.

Apesar do êxito obtido pelos adeptos de PbD, a aderência a tais técnicas não exaure o conteúdo do artigo 46 da LGPD, uma vez que esta abordagem é fundamentada em uma concepção de privacidade que é parte, mas não a totalidade, do conteúdo do direito à proteção de dados pessoais.[29] Em particular, a efetividade da revisão de uma decisão automatizada está menos ligada ao controle dos fluxos de dados a respeito de uma pessoa—que é o objeto primário de PbD—e mais a questões de transparência dos processos decisórios e da existência de informação suficiente a respeito de um sistema de tomada de decisão, que permitam ao titular

[29] A respeito da distinção conceitual entre privacidade e proteção de dados pessoais no ordenamento jurídico europeu, ver (KOKOTT; SOBOTTA, 2013). A fundamentação de tais direitos no ordenamento brasileiro também daria margem a similar distinção, mas o desenvolvimento desta excede os limites deste artigo.

de dados avaliar se deve ou não buscar a revisão da decisão a que foi submetido. E é possível que estes dois objetivos colidam em casos práticos (VEALE; BINNS; AUSLOOS, 2018). Consideremos o problema de identificar se um processo decisório é enviesado ou não. A busca por padrões de distorção ou discriminação em processos de tomada de decisão automatizada pode exigir um aumento na coleta de dados pessoais, para permitir que sejam identificados padrões sistêmicos que escapariam à análise de uma decisão considerada individualmente (VAN BEKKUM; BORGESIUS, 2023). Porém, um dos princípios que pautam PbD é a minimização dos dados pessoais em tratamento (LIMA; ALMADA; MARANHÃO, 2022, seç. 3.5), introduzindo assim tensões entre os requisitos de projeto que conduzem a um maior grau de privacidade e aqueles que conduzem a uma maior efetividade do direito à revisão (BAYAMLIOĞLU, 2022).

O conflito entre PbD e outros imperativos cobertos pela proteção de dados pessoais é um exemplo dos juízos de valor envolvidos no projeto de sistemas técnicos. Diante de conflitos entre valores juridicamente protegidos, a estratégia regulatória de proteção de dados *by design* acaba por delegar para os agentes de tratamento importantes decisões a respeito de como sopesar direitos fundamentais (ALMADA, 2023). Uma vez que este sopesamento já apresenta desafios substanciais para os tribunais,[30] os agentes de tratamento—em especial aqueles de pequeno e médio porte—necessitarão de diretrizes interpretativas para encontrar boas soluções para estes conflitos de direitos. Caso contrário, decisões de atores privados acabarão por causar mais dano do que elas buscam prevenir.

A proteção de dados *by design* é um instrumento que torna os agentes de tratamento garantidores da efetividade do sistema da LGPD. No contexto das decisões automatizadas, esta construção garante que os riscos decorrentes do uso de inteligência artificial e outras tecnologias de automação sejam mapeados pelos atores em melhor condição de resolvê-los. Porém, a diversidade de contextos de aplicação destas tecnologias e de valores em jogo nestes contextos faz com que a regulação baseada em risco produza enunciados normativos consideravelmente amplos, como examinado nesta seção. Assim, sua implementação exigirá a difusão de diretrizes e conhecimento que permitam o adequado mapeamento dos riscos existentes e soluções aplicáveis aos casos concretos.

[30] A respeito do sopesamento como problema jurídico, ver (SILVA, V. A., 2021, p. 120–122) e, no domínio da tecnologia, (GARBEN, 2020; SARTOR, 2016), dentre outros.

O PAPEL DA ANPD NA REGULAÇÃO DAS DECISÕES AUTOMATIZADAS

A regulação de decisões automatizadas no ordenamento brasileiro, como discutido *supra*, não é exaustiva em relação às preocupações geradas e sua abordagem traz espaço para importantes dúvidas interpretativas. Os dispositivos a respeito da proteção de dados *by design* fornecem diretrizes gerais para a avaliação dos riscos associados ao tratamento de dados pessoais, mas não indicam medidas específicas a serem adotadas ou mesmo riscos específicos ao contexto das decisões automatizadas. Já o artigo 20 da LGPD traz alguns exemplos de decisões automatizadas que o legislador julga criar riscos particularmente relevantes[31] e indica salvaguardas específicas para o cenário de automação: o direito à revisão (LGPD, Artigo 20, *caput*), o direito à explicação (LGPD, Artigo 20, § 1°)[32] e, nos casos em que este não se aplica por razões de segredo comercial ou industrial, a realização de auditorias para a verificação de potenciais aspectos discriminatórios (LGPD, Artigo 20, § 2°). Resta, contudo, uma série de problemas informacionais: como diagnosticar que outras medidas são necessárias para sistemas em concreto? Como identificar as medidas aplicáveis em contextos de decisões parcialmente automatizadas? O que se deve fazer a respeito dos efeitos sistêmicos da automação? Todas estas questões permanecem em aberto a partir da leitura do texto da lei, de forma que os agentes de tratamento de dados pessoais e os tribunais precisarão extrair respostas de outras fontes.

Uma fonte imediata para suprir esses aspectos na aplicação da LGPD às decisões automatizadas é a Autoridade Nacional de Proteção de Dados (ANPD). O Artigo 55-J da LGPD confere a esta autoridade diversas competências relevantes para a presente discussão, dentre elas encontram-se o poder de fiscalização e imposição de sanções em casos de tratamento em descumprimento à legislação (inciso IV), a difusão de conhecimento a respeito da proteção de dados pessoais (incisos V e VI), estimular a adoção de normas técnicas que facilitem o controle dos titulares sobre seus dados (inciso VII), e criar normas infralegais a respeito de vários aspectos da proteção de dados pessoais (incisos X, XIII e XVIII). Em particular, a autoridade é competente para estabelecer padrões técnicos mínimos que tornem aplicáveis as medidas de proteção de dados *by design*, considerando a natureza das informações tratadas, o estado atual da tecnologia e as características específicas do tratamento (LGPD, Artigo 46, § 1°). Ao fazê-

[31] O rol não-exaustivo do *caput* do artigo 20 da LGPD: "…incluídas as decisões destinadas a definir o seu perfil pessoal, profissional, de consumo e de crédito ou os aspectos de sua personalidade."
[32] Há ampla discussão na literatura a respeito do conteúdo das informações que devem ser providas neste contexto. Ver (MARANHÃO; COZMAN; ALMADA, 2021; MULHOLLAND; FRAJHOF, 2019; SÁ; LIMA, 2020).

lo, a ANPD pode criar normas que atendam às demandas específicas de determinadas operações de tratamento, bem como reconhecer e divulgar boas práticas adotadas por agentes de tratamento de dados (LGPD, Artigo 50, § 3º).

Dentro deste quadro, competiria à ANPD decretar que a revisão de decisões automatizadas seja, em certos casos, necessariamente feita por um humano. Como examinado anteriormente neste artigo, uma eventual regra nesse sentido não poderia ter caráter geral, uma vez que o processo legislativo da LGPD acabou por rejeitar uma regra geral de revisão por humanos. Tal rejeição, contudo, não veda a exigência de revisão humana em casos específicos. O artigo 46, § 1º, da LGPD autoriza a ANPD a dispor sobre padrões técnicos mínimos para proteger dados pessoais de tratamento ilícito ou inadequado. De acordo com esta autorização, as disposições da autoridade devem considerar "[...] a natureza das informações tratadas, as características específicas do tratamento e o estado atual da tecnologia, [...] assim como os princípios previstos no caput do art. 6º desta Lei". Uma vez que a tomada de decisão automatizada, como definida pela LGPD no artigo 20, é uma forma de tratamento de dados pessoais,[33] tal dispositivo permite que a autoridade de proteção de dados estabeleça requisitos técnicos mínimos que devem ser seguidos pelos particulares ao construir sistemas automatizados para tomada de decisão.

Em uma primeira análise, a remoção da previsão explícita de revisão por pessoa natural fornece um sinal de que tal regulamentação deve estimular o desenvolvimento de tecnologias que permitam uma revisão de decisões automatizadas por meios eles próprios automatizados, ou no mínimo ser aberta a esta possibilidade. De fato, a pesquisa acadêmica em engenharia de software e em inteligência artificial vem desenvolvendo técnicas que se prestam a tais propósitos, como a representação computacional de normas de proteção de dados (ROBALDO *et al.*, 2020), o uso de métricas de discriminação para detectar vieses em decisões algorítmicas (WACHTER; MITTELSTADT; RUSSELL, 2021) ou mesmo o uso de abordagens formais para a validação de software (KROLL *et al.*, 2017). Contudo, as abordagens que hoje existem são insuficientes para capturar elementos importantes para diagnosticar se houve ou não uma lesão a interesses juridicamente protegidos: por exemplo, sistemas computacionais encontram

[33] A menos que envolva apenas dados não-pessoais.

dificuldades para lidar com a ambiguidade inerente aos textos jurídicos[34] e com formas de comunicação como o humor, que são altamente dependentes de informação a respeito do contexto de comunicação (RAYZ; RASKIN, 2019). Ademais, as diversas formas de opacidade associadas aos sistemas computacionais (BURRELL, 2016) podem ser um obstáculo à legitimidade de sistemas automatizados de revisão: como ter certeza de que o sistema revisor está efetuando uma avaliação justa e não somente validando os *outputs* anteriores?

Estes e outros obstáculos fazem com que, ao menos no atual estado da arte, a revisão sem envolvimento humano seja viável apenas em tarefas particularmente simples, e não na maioria das situações em que se usa ou planeja usar uma ferramenta de decisão automatizada. Uma primeira contribuição que a ANPD poderia desempenhar seria clarificar as condições em que a revisão pode ser delegada. Tal clareza pode ser produzida através de mecanismos de ação direta da ANPD, como a edição de normas e diretrizes que estipulem casos em que a revisão automatizada é possível—ou ao menos forneçam critérios para que o controlador avalie a viabilidade desta. Além disso, a ANPD pode contribuir para a clareza do cenário regulatório através de mecanismos indutivos, como a difusão de conhecimento a respeito do estado da arte das decisões automatizadas ou o reconhecimento de práticas de autorregulação criadas no escopo do Artigo 50 da LGPD e que proponham boas diretrizes para avaliar a viabilidade de automação. Por estes mecanismos diretos e indiretos, a ANPD poderia trazer maior certeza em relação aos cenários em que a revisão puramente automatizada é possível sem prejuízo aos titulares de dados.

Quando a automação do processo revisional for impossível, o cumprimento da obrigação imposta ao controlador do tratamento de dados pessoais pelo *caput* do Artigo 20 da LGPD exigirá a adoção de medidas de revisão envolvendo pessoa natural. Nestes casos de impossibilidade, a ANPD estaria autorizada a estabelecer padrões técnicos que estipulem que a revisão seja feita por um humano, uma vez que a participação de um humano se torna uma condição necessária para o efetivo direito à revisão. Dada essa base, eventual padronização nesse sentido seria necessariamente transitória, perdendo sua base legal no caso de surgimento de tecnologias capazes de viabilizar a revisão automatizada, mas contribuiria para a segurança jurídica ao tornar explícito o que o melhor conhecimento técnico a respeito do tema reputa

[34] A respeito das abordagens para representar textos e raciocínios jurídicos em forma computacional, ver, e.g., (MARANHÃO; FLORÊNCIO; ALMADA, 2021, seç. 3–4)

possível ou impossível. Tendo em vista que o uso de sistemas de decisão automatizada não é (somente) uma elaboração da ficção científica, mas já faz parte do cotidiano de muitos brasileiros, a adequada definição de requisitos técnicos para a revisão de decisões automatizadas—incluindo talvez a exigência de um revisor humano—contribuirá para a melhor proteção dos direitos do titular de dados pessoais.

A ANPD pode, também, contribuir para a governança da IA ao exercer seus poderes de fiscalização. Um primeiro instrumento para este fim é o poder de realizar auditorias para verificar aspectos discriminatórios em decisões automatizadas (LGPD, art. 20, § 2º), que a ANPD pode exercer quando os agentes de tratamento de dados invocam segredos comerciais ou industriais para não revelar informações ao titular de dados. Para além desta hipótese potencialmente restrita de aplicação, a ANPD também tem o poder de requisitar relatórios de impacto à proteção de dados pessoais (LGPD, Artigos 10, § 3º, e 38, *caput*), cuja realização exige que os agentes de tratamento mapeiem os riscos envolvidos em uma operação de tratamento de dados. Tais relatórios e auditorias não só fornecem subsídios para futuras ações sancionadoras da autoridade—caso se mostrem necessárias nos termos da lei—mas também podem eles mesmos induzir ajustes no comportamento dos agentes sujeitos a estas atividades de prestação de contas (BUSUIOC, 2021).

O poder da ANPD em especificar os contornos dos dispositivos da LGPD não é ilimitado. Tendo em vista que o poder normativo de decretos e regulamentos é limitado à "fiel execução" do disposto na lei (CRFB, Artigo 84, IV), as normas emitidas pela autoridade não podem criar hipóteses de aplicação de dispositivos ou eliminar hipóteses previstas nestes. É questionável, por exemplo, se a ANPD tem competência para definir situações em que a revisão não seria obrigatória. Tal estipulação seria um mecanismo para filtrar a amplitude do conceito de "interesse" usado no *caput* do Artigo 20 da LGPD, desobrigando a revisão em casos julgados de menor impacto. Estabelecer um filtro nestas linhas, contudo, tornaria o direito à revisão inaplicável a situações em que o texto da LGPD prevê sua incidência, ao contrário da obrigação de revisão por humano, que meramente estipula o modo como o direito criado na lei deve ser assegurado. Para evitar que a obrigação da revisão gere um fardo excessivo para os agentes de tratamento, a ANPD pode lançar mão de outros mecanismos, como o estabelecimento de padrões simplificados para a revisão de decisões de caráter trivial ou de limitado impacto nos

interesses do titular de dados. Desta forma, a autoridade poderia tornar mais factível o direito à revisão sem, no entanto, exceder os limites de sua competência legal.

Ao enfrentar obstáculos como os elencados acima, a ANPD pode contribuir para que agentes de tratamento de dados pessoais possam diagnosticar e responder aos riscos associados a contextos de decisão automatizada. Por meio de sua produção normativa e técnica, esta autoridade pode atuar como difusora do conhecimento técnico existente a respeito das aplicações de inteligência artificial. Ao fazê-lo, a autoridade traria aportes para resolver disputas axiológicas como as descritas na seção anterior, bem como facilitaria o acesso de pequenas empresas e organizações do terceiro setor ao conhecimento necessário para o tratamento de dados pessoais. Neste papel, a atuação da ANPD coexiste com a atuação do setor privado, do terceiro setor, e de sistemas protetivos já existentes—como a ampla atuação do Ministério Público no campo da defesa do consumidor— no desenvolvimento de padrões técnicos para a proteção de dados (LGPD, Artigo 51), bem como no estabelecimento de regras de boas práticas e governança (LGPD, Artigo 40) e na criação de mecanismos de certificação.[35] Por último, mas não menos importante, a atuação sancionadora da ANPD, bem como as vindouras decisões dos tribunais a respeito da LGPD, servirão para verificar quais medidas técnicas, administrativas e de segurança se prestam à proteção de dados no caso concreto. Desta forma, o modelo brasileiro de proteção de dados pessoais oferece diversos mecanismos para sanar a imprecisão decorrente de sua ênfase na prevenção de riscos, ainda que seu funcionamento dependa da cooperação entre os diversos atores envolvidos.

CONSIDERAÇÕES FINAIS

Decisões automatizadas são uma questão saliente nas sociedades modernas, para a qual a LGPD propõe uma resposta dupla. De um lado, o direito à revisão das decisões automatizadas introduz um instrumento para que estes titulares de dados busquem seus interesses sem precisar recorrer à jurisdição administrativa ou judicial. De outro, a adoção de medidas técnicas e administrativas voltadas aos sistemas de tomada de decisão faz com que os agentes de tratamento de dados tenham de adotar uma postura proativa no diagnóstico e prevenção de potenciais danos oriundos das decisões automatizadas. Contudo, a diversidade de contextos de aplicação de sistemas de tomada de decisão automatizada, bem como a possibilidade de

[35] A LGPD, em seu Artigo 33, menciona a certificação como uma possibilidade no caso das transferências internacionais de dados.

automação parcial de decisões, faz com que tais normas devam cobrir uma vasta gama de casos e danos a prevenir.

Para assegurar a efetiva prevenção destes danos, a LGPD impõe diversos deveres aos agentes de tratamento de dados, de forma a obrigar tais agentes a avaliar e enfrentar os potenciais danos decorrentes do uso de decisões automatizadas. A alternativa seria a adoção de dispositivos detalhados para responder às minúcias de cada potencial caso de aplicação, uma solução que aumentaria a complexidade já elevada da lei e traria riscos de que ela se tornasse desatualizada diante das mudanças tecnológicas. Tal decisão, todavia, traz dificuldades para os agentes de tratamento de dados pessoais, que passam a precisar realizar uma avaliação contextual dos riscos decorrentes da automação de decisões, bem como adotar medidas que podem ter elevada complexidade técnica ou administrativa.

Sustentamos ao longo do artigo que este problema é fruto da novidade da lei e da complexidade dos cenários de uso das tecnologias de decisão automatizada. Portanto, a implementação do direito à revisão e das medidas de proteção de dados *by design* se beneficiará do conhecimento desenvolvido pelos vários atores sociais conforme a LGPD adquire maturidade. Para que isso ocorra, no entanto, faz-se necessária a ativa cooperação entre a ANPD e os atores públicos e privados que implementam sistemas de decisão automatizada. Caso contrário, os dispositivos que regulam decisões automatizadas correm o risco de se tornar uma ficção.

Mesmo com esta cooperação ativa, a legislação de proteção de dados é insuficiente para oferecer plena proteção diante dos riscos identificados nas seções anteriores. Na primeira seção, apresentamos alguns dos limites dos instrumentos centrados na noção de decisão automatizada. A seguir, apontamos que a regulação *by design* pode ser usada para estender o alcance da legislação de proteção de dados para cobrir outros riscos advindo do uso da IA, mas, ao fazê-lo, delega decisões a respeito de direitos fundamentais para os agentes de tratamento de dados. Por fim, vimos que a ANPD pode suprir balizas para estas decisões, mas está sujeita a certos limites ao fazê-lo, em especial no que tange ao seu poder normativo. As propostas formuladas nas seções anteriores não devem, portanto, ser vistas como uma alternativa à introdução de normas específicas para a IA, como aquelas atualmente sob debate no contexto do PL 21/2020 e a proposta de Substitutivo recentemente submetida pela Comissão de Juristas do Senado Federal. Em trabalhos futuros, pretendemos explorar em que medida o conteúdo do Substitutivo

supre as limitações aqui apontadas e dialoga adequadamente com a interpretação da LGPD aqui proposta.

Ainda assim, os instrumentos examinados ao longo do artigo trazem ferramentas versáteis que podem ser utilizadas para enfrentar diversos dos riscos que o uso da IA pode trazer aos titulares de dados, e muitos dos sistemas de maior impacto social dependem de dados pessoais para seu funcionamento. Desta forma, a proteção de dados pessoais é necessária—ainda que não suficiente—para a adequada regulação dos sistemas de IA e seu uso na tomada de decisão.

REFERÊNCIAS

ABREU, C. M. Discriminação de preços na economia digital: limites entre a eficiência e o abuso do poder econômico. **Revista do IBRAC**, *[S. l.]*, v. 24, n. 2, p. 309-329, 2018.

ALMADA, M. Regulation by design and the governance of technological futures. **European Journal of Risk Regulation**, *[S. l.]*, early access, 2023.

ARTICLE 29 WP. **Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679**. Bruxelas: Article 29 Working Party, 2018.

BAYAMLIOĞLU, E. The right to contest automated decisions under the General Data Protection Regulation: Beyond the so-called "right to explanation". **Regulation & Governance**, *[S. l.]*, v. 16, n. 4, p. 1058-1078, 2022.

BEIGANG, F. On the Advantages of Distinguishing Between Predictive and Al-locative Fairness in Algorithmic Decision-Making. **Minds and Machines**, *[S. l.]*, v. 32, n. 4, p. 655-682, 2022.

BIBAL, A. *et al.* Legal requirements on explainability in machine learning. **Artificial Intelligence and Law**, *[S. l.]*, v. 29, n. 2, p. 149-169, 2021.

BINNS, R.; VEALE, M. Is that your final decision? Multi-stage profiling, selective effects, and Article 22 of the GDPR. **International Data Privacy Law**, *[S. l.]*, v. 11, n. 4, 2021, p. 319-332.

BORDT, S. *et al.* Post-Hoc Explanations Fail to Achieve their Purpose in Adversarial Contexts. *In*: **Proceedings of FAccT '**22. Nova York: ACM, 2022, p. 891-905.

BRASIL. **Lei 13.709**, de 14 de agosto de 2018 (Lei Geral de Proteção de Dados Pessoais).

BRYSON, J. J.; DIAMANTIS, M. E.; GRANT, T. D. Of, for, and by the people: the legal lacuna of synthetic persons. **Artificial Intelligence and Law**, *[S. l.]*, v. 25, n. 3, p. 273-291, 2017.

BURRELL, J. How the machine 'thinks': Understanding opacity in machine learning algorithms. **Big Data & Society**, *[S. l.]*, v. 3, n. 1, p. 1-12, 2016.

BUSUIOC, M. Accountable Artificial Intelligence: Holding Algorithms to Account. **Public Administration Review**, *[S. l.]*, v. 81, n. 5, p. 825-836, 2021.

BUSUIOC, M.; CURTIN, D.; ALMADA, M. Reclaiming transparency: contesting the logics of secrecy within the AI Act. **European Law Open**, *[S. l.]*, v. 2, n. 1, p. 79-105, 2023.

BYGRAVE, L. A. Article 25. Data protection by design and by default. In: KUNER, C.; BYGRAVE, L. A; DOCKSEY, C. (Org.). **The EU General Data Protection Regulation (GDPR)**: A Commentary. Oxford: Oxford University Press, 2020. p. 571-581.

BYGRAVE, L. A. Data Protection by Design and by Default: Deciphering the EU's Legislative Requirements. **Oslo Law Review**, Oslo, v. 4, n. 2, p. 105-120, 2017.

CITRON, D. K.; PASQUALE, F. The Scored Society: Due Process for Automated Predictions. **Washington Law Review**, Seattle, v. 89, p. 2-33, 2014.

COBBE, J.; SINGH, J. Regulating Recommending: Motivations, Considerations, and Principles. **European Journal of Law and Technology**, *[S. l.]*, v. 10, n. 3, 2019.

COBBE, J.; VEALE, Michael; SINGH, Jatinder. Understanding accountability in algorithmic supply chains. *In*: **Proceedings of FAccT 2023**. Nova York: ACM, 2023, p. 1186-1197.

DIETTRICH, L.; DE SOUZA, F.; GUERREIRO, A. Development of credit scores with telco data using Machine Learning and agile methodology in Brazil. In: **SAS Global Forum 2020**, *[S.l.]*: SAS, 16 jun. 2020.

EDPB. **Recommendations 01/2020 on measures that supplement transfer tools to ensure compliance with the EU level of protection of personal data**. European Data Protection Board, 10 de novembro de 2020.

FINCK, M.; PALLAS, F. They who must not be identified—distinguishing personal from non-personal data under the GDPR. **International Data Privacy Law**, *[S. l.]*, v. 10, n. 1, p. 11–36, 2020.

GARBEN, S. Fundamental rights in EU copyright harmonization: Balancing without a solid framework: Funke Medien, Pelham, Spiegel Online. **Common Market Law Review**, Leiden, v. 57, n. 6, p. 1909-1932, 2020.

GELLERT, R. Understanding the notion of risk in the General Data Protection Regulation. **Computer Law & Security Review**, *[S. l.]*, v. 34, n. 2, p. 279-288, 2018.

GORDON, G.; RIEDER, B.; SILENO, G.. On mapping values in AI Governance. **Computer Law & Security Review**, *[S. l.]*, v. 46, 2022.

HILDEBRANDT, M. Privacy as Protection of the Incomputable Self: From Agnostic to Agonistic Machine Learning. **Theoretical Inquiries in Law**, *[S. l.]*, v. 20, n. 1, p. 83-121, 2019.

HOHFELD, W. N. Some Fundamental Legal Conceptions as Applied in Judicial Reasoning. **The Yale Law Journal**, New Haven, v. 23, n. 1, p. 16-59, 1913.

HOSNI, D. S. S.; MARTINS, P. B. L. Tomada de Decisão Automatizada e a Regulamentação da Proteção de Dados: Alternativas Coletivas Oferecidas pela Lei Geral de Proteção de Dados. **Internet & Sociedade**, São Paulo, v. 1, n. 2, p. 77-101, 2020.

KOKOTT, J.; SOBOTTA, C. The distinction between privacy and data protection in the jurisprudence of the CJEU and the ECtHR. **International Data Privacy Law**, *[S. l.]*, v. 3, n. 4, p. 222-228, 2013.

KROLL, J. *et al.* Accountable Algorithms. *University of Pennsylvania Law Review*, Filadélfia, v. 165, n. 3, pp. 633-705, 2017.

LAGIOIA, F.; SARTOR, G. Artificial intelligence in the big data era: risks and opportunities. In: CANNATACCI, J.; FALCE, V.; POLLICINO, O. (Org.). **Legal Challenges of Big Data**. Northampton: Edward Elgar, 2020. p. 280-307.

LIMA, C. C. C.; ALMADA, M.; MARANHÃO, J. Data protection by design e data protection by default. Visão teórica e prática à luz da LGPD e do GDPR. In: VAINZOF, R.; SERAFINO, D.; STEINWASCHER, A. (Org.). **Legal Innovation**: O Futuro do Direito e o Direito do Futuro. São Paulo: Thomson Reuters Brasil, 2022. p. 265-288.

LÓPEZ, Nuria. Um direito, um dever: guia para o art. 20 da LGPD. In: OPICE BLUM, Renato (Org.). *Proteção de Dados - Desafios e Soluções na Adequação à Lei*. Rio de Janeiro: Forense, 2020.

MALGIERI, G.; COMANDÉ, G. Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation. **International Data Privacy Law**, *[S. l.]*, v. 7, n. 4, p. 243-265, 2017.

MARANHÃO, J.; ALMADA, M. Inteligência Artificial no Setor de Saúde: Ética e Proteção de Dados. In: DALLARI, A. B.; MONACO, G. F. C. (Org.). *LGPD na Saúde*. 1ª edição ed. São Paulo: Thomson Reuters Brasil, 2021. p. 357-372.

MARANHÃO, J.; CAMPOS, R. R. Proteção de Dados de Crédito na Lei Geral de Proteção de Dados. **Direito Público**, *[S. l.]*, v. 16, n. 90, p. 132-154, 2019.

MARANHÃO, J.; COZMAN, F. G.; ALMADA, M. Concepções de explicação e do direito à explicação de decisões automatizadas. In: VAINZOF, R.; GUTIERREZ, A. (Org.). **Inteligência Artificial**: Sociedade, Economia e Estado. São Paulo: Thomson Reuters Brasil, 2021. p. 137-154.

MARANHÃO, J.; FLORÊNCIO, J. A.; ALMADA, M. Inteligência artificial aplicada ao direito e o direito da inteligência artificial. **Suprema** - Revista de Estudos Constitucionais, Brasília, v. 1, n. 1, p. 154-180, 2021.

MARRAFON, M. A.; COUTINHO, L. L. C. L. Princípio da Privacidade por Design: Fundamentos e Efetividade Regulatória na Garantia do Direito à Proteção de Dados. **Revista Eletrônica Direito e Política**, *[S. l.]*, v. 15, n. 3, p. 955-984, 2020.

MENDES, Gilmar Ferreira. **Referendo na Medida Cautelar na Ação Direta de Inconstitucionalidade 6.389 Distrito Federal**. Brasília: Supremo Tribunal Federal, 7 de maio de 2020

MULHOLLAND, C.; FRAJHOF, I. Z. Inteligência artificial e a lei de proteção de dados pessoais: breves anotações sobre o direito à explicação perante a tomada de decisões por meio de machine learning. In: FRAZÃO, A.; MULHOLLAND, C. (Org.). **Inteligência Artificial e Direito**: Ética, Regulação e Responsabilidade. São Paulo: Thomson Reuters Brasil, 2019.

PRESIDÊNCIA DA REPÚBLICA. **Mensagem n° 288, de 8 de julho de 2019**. Mensagem de veto presidencial. Brasília: Presidência da República, 2019.

RAJI, I. D. *et al.* The Fallacy of AI Functionality. *In:* **Proceedings of FAccT '22**. New York: Association for Computing Machinery, 2022. p. 959-972.

RAYZ, J. T.; RASKIN, V. Fuzziness and Humor: Aspects of Interaction and Computation. **Advances in Intelligent Systems and Computing**. Cham: Springer, 2019. p. 655-666.

RECHTBANK AMSTERDAM. **Decisão no caso HA RK 20-207**. Amsterdam: Rechtbank Amsterdam, 11 mar. 2021.

RIVELLI, F.; SILVEIRA, R. F. Privacy By Design E Privacy By Default-Proteção da privacidade na área da saúde desde a concepção (e por padrão). In: DALLARI, A. B.; MONACO, G. F. C. (Org.). **LGPD na Saúde**. 1ª edição ed. São Paulo: Thomson Reuters Brasil, 2021. p. 237-247.

ROBALDO, L. *et al.* Formalizing GDPR Provisions in Reified I/O Logic: The DAPRECO Knowledge Base. **Journal of Logic, Language and Information**, *[S. l.]*, v. 29, n. 4, p. 401-449, 2020.

ROBERTO, E. Responsabilidade civil pelo uso de sistemas de inteligência artificial: em busca de um novo paradigma. **Internet & Sociedade**, São Paulo, v. 1, n. 1, p. 121–143, 2020.

SÁ, M. F. F. De; LIMA, T. M. M. Inteligência artificial e Lei Geral de Proteção de Dados Pessoais: o direito à explicação nas decisões automatizadas. **Revista Brasileira de Direito Civil - RBDCivil**, Rio de Janeiro, v. 26, n. 04, p. 227, 2020.

SARRA, C. Put Dialectics into the Machine: Protection against Automatic-decision-making through a Deeper Understanding of Contestability by Design. **Global Jurist**, *[S. l.]*, v. 20, n. 3, 2020.

SARTOR, G. The right to be forgotten: balancing interests in the flux of time. **International Journal of Law and Information Technology**, *[S. l.]*, v. 24, n. 1, p. 72-98, 1 mar. 2016.

SCHAFER, B. Death by a Thousand Cuts: Cumulative Data Effects and the Corbyn Affair. **Datenschutz und Datensicherheit - DuD**, *[S. l.]*, v. 45, n. 6, p. 385-390, 1 jun. 2021.

SILVA, T. Visão computacional e racismo algorítmico: branquitude e opacidade no aprendizado de máquina. **Revista da ABPN**, *[S. l.]*, v. 12, n. 31, 2020.

SILVA, V. A. **Direito Constitucional Brasileiro**. São Paulo: Edusp, 2021.

SILVA, P.; MEDEIROS, J. **A polêmica da revisão (humana) sobre decisões automatizadas.** Rio de Janeiro: ITS Rio, 10 dez. 2019. Disponível em https://feed.itsrio.org/a-pol%C3%AAmica-da-revis%C3%A3o-humana-sobre-decis%C3%B5es-automatizadas-a81592886345. Acesso em: 16 set. 2021.

TAMÒ-LARRIEUX, A. **Designing for Privacy and its Legal Framework**: Data Protection by Design and Default for the Internet of Things. Cham: Springer, 2018.

TOBBIN, R. A.; CARDIN, V. S. G. Perfis Informacionais e Publicidade Comportamental: Direito à Autodeterminação Informativa e a Proteção de Dados Pessoais no Ambiente Virtual. **Anais do Congresso Brasileiro de Processo Coletivo e Cidadania**, Ribeirão Preto, n. 8, p. 1260-1276, 12 dez. 2020.

TOMASSETTI, J. Algorithmic Management, Employment, and the Self in Gig Work. In: DAS ACEVEDO, D. (Org.). **Beyond the Algorithm**: Qualitative Insights for Gig Work Regulation. Cambridge: Cambridge University Press, 2020. p. 123-145.

TRIBUNAL EUROPEU DE DIREITOS HUMANOS. **I v. Finland (application 20511/03)**. Estrasburgo: Tribunal Europeu de Direitos Humanos, 17 jul. 2008.

UNIÃO EUROPEIA. **Regulamento (UE) 2016/679** do Parlamento Europeu e do Conselho de 27 de abril de 2016 relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados e que revoga a Diretiva 95/46/CE (Regulamento Geral sobre a Proteção de Dados) (Texto relevante para efeitos do EEE).

VALE, S. B.; ZANFIR-FORTUNA, G. **Automated Decision-Making Under the GDPR**: Practical Cases from Courts and Data Protection Authorities. Bruxelas: Future of Privacy Forum, 2022.

VAN BEKKUM, M.; BORGESIUS, F. Z. Using sensitive data to prevent discrimination by artificial intelligence: Does the GDPR need a new exception? **Computer Law & Security Review**, *[S. l.]*, v. 48, 2023.

VEALE, M.; BINNS, R.; AUSLOOS, J. When data protection by design and data subject rights clash. **International Data Privacy Law**, *[S. l.]*, v. 8, n. 2, p. 105-123, 2018.

WASHINGTON, A. L. How to argue with an algorithm: Lessons from the Compas-ProPublica debate. **Colorado Technology Law Journal**, Boulder, v. 17, n. 1, p. 131-160, 2018.

WACHTER, S.; MITTELSTADT, B.; RUSSELL, C. Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI. **Computer Law & Security Review**, *[S. l.]*, v. 41, 2021.

WINFIELD, A. F. T. *et al.* IEEE P7001: A Proposed Standard on Transparency. **Frontiers in Robotics and AI**, *[S. l.]*, v. 8, 2021.

WINNER, L. Do Artifacts Have Politics? **Daedalus**, Cambridge (Massachusetts), v. 109, n. 1, p. 121-136, 1980.

Sobre os(as) autores(as):

Marco Almada | *E-mail: Marco.Almada@eui.eu*

Marco Almada é pesquisador no Instituto Universitário Europeu (EUI), onde realiza pesquisas de doutorado na área de regulação da inteligência artificial. Possui formação em direito (mestre pelo EUI; bacharel pela USP) e computação (mestre e bacharel pela Unicamp).

Juliano Maranhão | *E-mail: julianomaranhao@usp.br*

Juliano Maranhão é professor associado da Faculdade de Direito da Universidade de São Paulo. Doutor em Direito pela Faculdade de Direito da USP, realizou pós-doutorado no Departamento de Ciência da Computação da Universidade de Utrecht, na Holanda (2005). Presidente do Instituto Lawgorithm de Inteligência Artificial, membro do Comitê Diretor da Associação Internacional de Inteligência Artificial e Direito e diretor do Instituto Legal Grounds. Foi assessor da presidência do CADE.

Data de submissão: 30 de maio de 2023.

Data do aceite: 19 de julho de 2023.