RENATO LEITE MONTEIRO

Desafios para a efetivação do direito à explicação na Lei Geral de Proteção de Dados do Brasil

Tese de Doutorado

Orientador: Professor Dr. Rafael Mafei Rabelo Queiroz

UNIVERSIDADE DE SÃO PAULO

FACULDADE DE DIREITO

SÃO PAULO

2021

RENATO LEITE MONTEIRO

Desafios para a efetivação do direito à explicação na Lei Geral de Proteção de Dados do Brasil

Tese de Doutorado apresentada à Banca Examinadora do programa de Pós-Graduação em Direito, da Faculdade de Direito da Universidade de São Paulo, na área de concentração Filosofia e Teoria Geral do Direito, sob a orientação do Prof. Dr. Rafael Mafei Rabelo Queiroz.

UNIVERSIDADE DE SÃO PAULO

FACULDADE DE DIREITO

SÃO PAULO

2021

Catalogação da Publicação
Serviço de Biblioteca e Documentação
Faculdade de Direito da Universidade de São Paulo

Monteiro, Renato Leite

Desafios para a efetivação do direito à explicação na Lei Geral de Proteção de Dados do Brasil; Renato Leite Monteiro; orientador Rafael Mafei Rabelo Queiroz - São Paulo, 2021.

Tese (Doutorado — Programa de Pós-Graduação em Filosofia do Direito e Teoria Geral do Direito) — Faculdade de Direito, Universidade de São Paulo, 2021.

1. Decisões automatizadas. 2. Dados pessoais. 3. Regulação. 4. Governança. 5. LGPD. I. Queiroz, Rafael Mafei Rabelo, orient. II. Título.

MONTEIRO, Renato Leite. *Desafios para a efetivação do direito à explicação na Lei Geral de Proteção de Dados do Brasil*. 2021. 385 f. Tese (Doutorado em Direito) – Faculdade de Direito, Universidade de São Paulo, São Paulo, 2021.

Tese de Doutorado apresentada à Banca Examinadora do programa de Pós-Graduação em Direito, da Faculdade de Direito da Universidade de São Paulo, na área de concentração Filosofia e Teoria Geral do Direito, sob a orientação do Prof. Dr. Rafael Mafei Rabelo Queiroz.

Aprovado em:

BANCA EXAMINADORA

Prof(a). Dr(a). _____ Instituição _____
Julgamento _____ Assinatura _____

Prof(a). Dr(a). _____ Instituição _____
Julgamento _____ Assinatura _____

Prof(a). Dr(a). _____ Instituição _____
Julgamento _____ Assinatura _____

Prof(a). Dr(a). _____ Instituição _____
Julgamento _____ Assinatura _____

Prof(a). Dr(a). _____ Instituição _____
Julgamento _____ Assinatura _____

O mund dá muitas voltas...

AGRADECIMENTOS

Sendo bem franco, a seção de agradecimentos foi uma das partes mais difíceis de ser escrita desta tese. A dificuldade é oriunda, justamente, da multiplicidade de pessoas que fizeram parte da minha vida e que a levaram até esse momento. Como se você fosse um produto de todas elas, e não somente por aquelas que dela fizeram parte durante o árduo período do doutoramento. Da mesma forma que somente é possível entender como chegamos até aqui quando paramos para olhar para trás e começamos a ligar os pontos das pessoas, lugares e experiências que nos formaram, que moldaram nosso senso crítico e nos fizeram quem somos. Por isso é tão difícil escrever agradecimentos. É fácil agradecer por atitudes pontuais, mas não por conjuntos de experiências, conhecimento acumulado e formação de personalidade, elementos que têm uma contribuição direta à forma como a presente pesquisa foi desenvolvida. Mas sim, há atores e atrizes que têm uma contribuição maior e específica e que merecem agradecimentos individualizados.

Primeiro, gostaria de agradecer imensamente ao meu orientador, o Prof. Rafael Mafei, pela sua paciência, compreensão e suporte ao longo de todos esses anos de doutoramento. Sei que muitas vezes não atendi às suas expectativas. Não foi nada fácil conseguir conciliar estudos e pesquisa com uma intensa rotina de trabalho, que sofreu várias reviravoltas, que muitas vezes consumiram precioso tempo que deveria ter sido dedicado ao objeto do presente estudo. Por isso, serei eternamente grato.

Agradeço também aos demais professores da Universidade de São Paulo, principalmente da Faculdade de Direito do Largo São Francisco, que proporcionaram diálogos propositivos de interseção entre as diferentes áreas do direito e os seus reflexos na proteção de dados. Ainda, aos membros da minha banca da qualificação, o Prof. Danilo Doneda, por quem nutro um enorme carinho e admiração. Sem o Prof. Danilo Doneda provavelmente eu, e toda uma geração de pesquisadores na área de proteção de dados, não existiria. Seu legado será sentido por décadas a fio e é de um valor inestimável. Agradeço também ao Prof.

Dennys Antonialli, por quem também nutro um respeito enorme. Seus apontamentos foram essenciais para o produto final dessa pesquisa.

Ao meu porto seguro, minha família. Meus pais, irmãs, sobrinhos, tios e tias, primos, primas, avôs e avós. Somente me foi possível realizar minhas andanças mundo afora por saber que sempre teria um lugar seguro para voltar. Sempre havia uma rede de suporte na qual eu poderia me segurar. Um especial agradecimento às mulheres dessa família, que são uma fonte inesgotável de admiração e respeito. Sua resiliência deveria servir de inspiração para toda uma geração de homens e mulheres que por vezes se encontram inseguros e na ausência de figuras fortes e íntegras a quem se espelharem. Vocês são exemplos a serem seguidos por nossos filhos, sobrinhos e netos ainda por vir.

Acrescento à família um irmão que a vida me deu, Lucas Tasquetto. Não o poderia colocar em outro “grupo”. Por vezes brincamos que somos o relacionamento sério mais longo de nossas vidas, mas desde o momento em que nos conhecemos a nossa parceria proporcionou alguns dos momentos mais importantes da minha vida. Esta tese com certeza não teria surgido sem o seu eterno apoio, irmandade e cumplicidade. E, é claro, à Mari, por estar sempre ao meu lado... Por isso e por muito mais, o agradeço e agradeço a todos que fazem parte dessa família linda.

Não consigo enumerar ou listar todos os amigos e amigas que gostaria, e que precisaria, agradecer. Sempre fui uma pessoa iluminada por amigos-irmãos únicos que, cada um em sua medida, contribuíram para a pessoa que hoje sou. Espalhados literalmente por todos os cantos do mundo, vocês estiveram presentes nos momentos mais felizes, nos mais tristes e nos mais desafiadores da minha vida. Hoje sou um pedaço de cada um de vocês e seria injusto com todos vocês se apenas os agradecesse. Vocês ajudaram a conectar os pontos de uma vida de aventuras, conhecimento e crescimento. Eu devo a vocês quem hoje eu sou.

Todavia, no mundo da proteção de dados, eu não poderia deixar de agradecer individualmente a duas pessoas, meu amigo-irmão Bruno Bioni e a minha grande amiga Sophie Kwasny. Bruno, ou Artilheiro, vocativo pelo qual nos

chamamos, entrou na minha vida por meio do interesse mútuo na proteção de dados. Dessa amizade surgiu essa tese, surgiu o Data Privacy Brasil, surgiu uma confiança e admiração sem igual. Seu caráter, seriedade, compreensão, irmandade e parceria são fontes diárias de admiração e respeito. Essa tese também é sua. Sophie, que abriu as portas para o início da minha atuação internacional na área de proteção de dados, se tornou uma grande amiga que me proporciona sorrisos sinceros ao mero relembrar de nossos momentos juntos. O que você fez pela nossa área é inigualável. Sua incansável luta pelo respeito a valores universais de direitos humanos é algo que sempre renova minha luta por propósitos quando me encontro perdido em um mundo apático, líquido e aparentemente amoral. Um verdadeiro alicerce que serve de sustentáculo e esperança para um mundo onde podemos, e devemos, tentar fazer a diferença. Muito obrigado!

Ao Data Privacy Brasil, instituição que surgiu dos devaneios de dois amigos que achavam que podiam contribuir para a educação, capacitação, pesquisa e para o desenvolvimento de uma cultura brasileira de privacidade e proteção de dados, e se tornou um dos principais centros de referência sobre o tema no Brasil e no mundo. Todos os dias aprendo com seus colaboradores incríveis, dedicados, brilhantes e humildes. Os frutos dessa dedicação já são colhidos país afora, seja na formação de toda uma nova geração de profissionais e pesquisadores na área de proteção de dados, seja na árdua e contínua luta contra práticas abusivas, desiguais e desproporcionais que podem impactar direitos fundamentais de milhares de pessoas. Acredito que vocês ainda não têm noção do impacto e da diferença que fazem na vida dessas pessoas. Por isso, e por muito mais, agradeço de forma imensurável.

Entretanto, dentre os colaboradores do Data Privacy Brasil, não poderia deixar de agradecer em especial aos pesquisadores Sinuhe Cruz, Leôncio Júnior, Caio Machado e Mariana Rielli, que colaboram ativamente com as discussões que resultaram nesta tese de doutorado. Aprendi com vocês muito mais do que vocês podem imaginar. Por isso, agradeço muito.

Ao Baptista Luz Advogados e, principalmente, ao seu time de proteção de dados e ao meu grande amigo Pedro Ramos. Vocês acreditaram em mim e juntos

criamos do zero um grupo de profissionais e uma área até hoje inigualável no Brasil. A dedicação, a resiliência e a paciência de vocês ajudaram a forjar uma equipe coesa, amiga e brilhante com quem aprendi imensamente e passei alguns dos melhores anos da minha vida. Um dia iremos todos juntos para Jeri, ainda irei cumprir essa promessa!

Por último, não poderia deixar de agradecer ao Twitter e ao nosso time global de Proteção de Dados. Vocês surgiram num *plot twist* da minha vida que nem mesmo o mais kafkiano dos roteiristas poderia prever e me permitiram ver a proteção de dados na prática de uma forma que nem mesmo os melhores livros conseguiriam descrever. A compreensão de todos vocês para que eu pudesse ao mesmo tempo me dedicar ao trabalho, ao Data Privacy Brasil e ao desenvolvimento desta tese permitiram a sua finalização. Vocês fazem a diferença em um mundo que diariamente nos testa, por vezes nos decepciona, mas que ressurge frequentemente melhor e mais forte.

Sei que provavelmente deixei muitas pessoas de fora, mas todos e todas estão comigo e estão em cada linha desta tese. Ela é o resultado de cada palavra, experiência e momento que vocês proporcionaram. Serei eternamente grato a todos vocês.

RESUMO

MONTEIRO, Renato Leite. *Desafios para a efetivação do direito à explicação na Lei Geral de Proteção de Dados do Brasil*. 2021. 385 p. Tese (Doutorado em Direito) – Faculdade de Direito, Universidade de São Paulo, São Paulo, 2021.

Decisões automatizadas cada vez mais controlam nossas vidas, gerenciadas por algoritmos, cujos resultados podem ter um impacto significativo sobre os cidadãos. Todavia, a maior presença dessas decisões no cotidiano é acompanhada de pouca transparência com relação ao seu funcionamento – o que torna mais complexa a identificação de práticas abusivas, discriminatórias ou, ainda, monopolísticas, que podem causar impactos nos planos individual e coletivo. Para mitigar tais efeitos, legislações nacionais e internacionais de proteção de dados tentam assegurar os direitos à transparência, à explicação e ao não estar sujeito a decisões automatizadas. A presente pesquisa realizou uma análise dos aspectos jurídicos da proteção de dados pessoais no Brasil e analisou, mais especificamente, a existência de um direito à explicação no contexto de decisões automatizadas, assim como os desafios acerca da sua implementação e execução. A principal hipótese desta pesquisa é que existe um direito à explicação no contexto de decisões automatizadas orientadas por algoritmos. Todavia, ainda é incerto como, na prática, instrumentalizar tal direito, levando em consideração: (i) a complexidade de sistemas algorítmicos, quase que opacos por natureza, principalmente nos que se valem de aprendizado de máquina para tomar suas decisões; (ii) os limites impostos pela própria legislação, como segredo de negócio e propriedade intelectual; e (iii) limitações cognitivas podem dificultar a compreensão de informações fornecidas. O objetivo deste trabalho é propor e colaborar com o desenvolvimento de elementos, instrumentos e critérios sob um viés técnico-jurídico que possam colaborar para explicações efetivas e úteis que permitam coibir práticas discriminatórias, abusivas e desproporcionais, nos planos individual e coletivo.

Palavras-chave: decisões automatizadas; algoritmos; dados pessoais; regulação; explicação; governança.

ABSTRACT

MONTEIRO, Renato Leite. *Challenges for the effectiveness of the right to explanation in the General Data Protection Law of Brazil*. 2021. 385 p. Thesis (Doctorate in Law) – Faculty of Law, University of São Paulo, São Paulo. 2021.

Automated decisions increasingly control our lives, managed by algorithms, whose results can have a significant impact on citizens. However, the increased presence of these decisions in everyday life is accompanied by little transparency regarding their operation - which makes it more complex to identify abusive, discriminatory or even monopolistic practices, which can have impacts on the individual and collective levels. To mitigate these effects, national and international data protection legislation attempts to ensure the rights to transparency, explanation, and not being subject to automated decisions. This research conducted an analysis of the legal aspects of protection of personal data in Brazil and analyzed, more specifically, the existence of a right to explanation in the context of automated decisions, as well as the challenges about its implementation and enforcement. The main hypothesis of this research is that a right to explanation exists in the context of automated algorithm-driven decisions. However, it is still uncertain how, in practice, to instrumentalize such a right, taking into consideration: (i) the complexity of algorithmic systems, almost opaque by nature, especially in those that rely on machine learning to make their decisions; (ii) the limits imposed by the legislation itself, such as trade secrets and intellectual property; and (iii) cognitive limitations may hinder the understanding of information provided. The objective of this paper is to propose and collaborate with the development of elements, instruments, and criteria from a technical-legal standpoint that may contribute to effective and useful explanations to curb discriminatory, abusive, and disproportionate practices at the individual and collective levels.

Keywords: automated decisions; algorithms; personal data; regulation; explanation governance.

RÉSUMÉ

MONTEIRO, Renato Leite. *Défis pour l'efficacité du droit à l'explication dans la loi générale sur la protection des données du Brésil.* 2021. 385 p. Thèse (Doctorat en Droit) – Faculté de Droit, Université de São Paulo, São Paulo. 2021.

Des décisions automatisées contrôlent de plus en plus nos vies. Gérées par les algorithmes, elles peuvent avoir une influence significative sur les citoyens. Pourtant, la majeure partie de ces décisions dans le quotidien manquent de transparence par rapport à son fonctionnement - ce qui rend plus complexe l'identification de pratiques abusives, discriminatoires ou, encore, monopolistique, ce qui peut causer des impacts tant au niveau individuel que collectif. Afin d'atténuer ces effets, des législations nationales et internationales de protection des données essayent d'assurer les droits à la transparence, à l'explication et au non-assujettissement des décisions automatisées. Cette étude a effectué une analyse des aspects juridiques de la protection des données personnelles au Brésil et a analysé plus particulièrement l'existence d'un droit à l'explication en ce qui concerne les décisions automatisées, ainsi que les défis posés par son implémentation et son exécution. L'hypothèse principale de cette recherche, c'est qu'il existe en effet un droit à l'explication lors des décisions automatisées orientées par des algorithmes. Il n'est cependant pas encore évident comment l'instrumentaliser, lorsqu'on considère : (i) la complexité des systèmes algorithmiques, presque opaques par nature, surtout ceux qui s'appuient sur l'apprentissage automatique pour prendre leurs décisions ; (ii) les limites imposées par la législation elle-même, comme les secrets d'affaire et la propriété intellectuelle ; (iii) des fonctions cognitives limitées qui peuvent faire obstacle à la compréhension des informations fournies. Par le moyen d'une perspective technique et juridique, le but de cette étude est donc de proposer des suggestions et de collaborer avec le développement d'éléments, d'instruments et de critères qui puissent faciliter des explications efficaces et utiles contre les pratiques discriminatoires, abusives et disproportionnées tant au niveau individuel que collectif.

Mots-clés: décisions automatisées; algorithmes; données personnelles; réglementation; explication; gouvernance.

LISTA DE GRÁFICOS

LISTA DE FIGURA E QUADROS

SUMÁRIO

INTRODUÇÃO

Nossas vidas são controladas por algoritmos. Eles estão presentes na definição da melhor rota para fugir do trânsito[1], na seleção de candidatos para vagas de trabalho[2], na determinação de penas de condenados por crimes[3], podem influenciar pleitos eleitorais[4] e na formulação de políticas públicas[5]. Todavia, uma grande parte desses algoritmos padece de uma grave opacidade, uma falta de transparência. Essa opacidade impede que as pessoas, e a sociedade, entendam e verifiquem se seus dados pessoais são tratados de forma legítima, adequada e proporcional[6], além da verificação se há aspectos discriminatórios ou que possam impactar desproporcionalmente direitos e liberdades fundamentais. Mitigar tal opacidade por meio de obrigações de transparência é um dos objetivos de leis que versam sobre o tratamento adequado de dados pessoais.

Regular o uso e o tratamento de dados pessoais é o principal objeto de leis de proteção de dados. Estas visam não somente proteger a privacidade, mas também garantir direitos fundamentais e liberdades individuais[7], que somente podem ser exercidos na sua completude caso seja garantido o uso adequado dos dados pessoais, entendidos como uma representação do indivíduo. Pode-se entender as leis de proteção de dados como um plexo regulatório que acaba por proteger outros

---

[1] BOEGLIN, Jack. The Costs of Self-Driving Cars: Reconciling Freedom and Privacy with Tort Liability in Autonomous Vehicle Regulation. p. 35, 2015.

[2] PURAM, K.; SADAGOPAL, G. US Patent 6,289,340 (2001). Consultant matching system and method for selecting candidates from a candidate pool by adjusting skill values. Disponível em: https://patents.google.com/patent/US6289340B1/en. Acesso em: 02 jul. 2021.

[3] ANGWIN, J. et al. (2016). “Machine Bias”. ProPublica. Disponível em: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. Acesso em: 02 jul 2021.

[4] INFORMATION COMMISSIONER’S OFFICE (2018). “Democracy disrupted? Personal information and political influence”. ICO. Disponível em: https://ico.org.uk/media/2259369/democracydisrupted-110718.pdf. Acesso em: 02 jul. 2021.

[5] NEMITZ, Paul. Constitutional Democracy and Technology in the Age of Artificial Intelligence. SSRN Scholarly Paper, no ID 3234336. Rochester, NY: Social Science Research Network, 18 ago. 2018. Disponível em: https://papers.ssrn.com/abstract=3234336. Acesso em: 02 jul. 2021.

[6] BRUNDAGE, Miles et al. The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation. arXiv:1802.07228 [cs], arXiv: 1802.07228, 20 fev. 2018. Disponível em: http://arxiv.org/abs/1802.07228. Acesso em: 27 maio 2019.

[7] WATCHER, Sandra. Show Me Your Data and I’ll Tell You Who You Are. Oxford Internet Institute London Lecture. 2018. Disponível em: https://www.oii.ox.ac.uk/videos/oii-london-lecture-show-me-your-data-and-ill-tell-you-who-you-are/. Acesso em: 02 jul. 2021.

direitos.[8] Os contextos nos quais decisões automatizadas têm impactado no exercício e acesso a uma série de direitos fundamentais são variados e complexos.[9] Todavia, a opacidade com a qual os dados pessoais são tratados impede que seus titulares tenham total compreensão de como suas vidas são impactadas.[10]

Tais contextos dão origem à necessidade de um novo construto normativo capaz de garantir proteção efetiva a seus titulares. É a partir desta demanda que se entende o chamado *direito à explicação*. Este pode ser entendido como o direito a receber informações suficientes e inteligíveis que permitam ao titular dos dados e à sociedade entenderem e compreenderem a lógica, a forma e os critérios utilizados para tratar dados pessoais e prever os seus impactos[11], com o fim de evitar práticas discriminatórias, ilegítimas e indesejadas, que podem ter impacto no plano individual e coletivo.

O direito à explicação seria regulado, em certa extensão, pela Lei Geral de Proteção de Dados do Brasil (Lei 13.709/2018, também chamada de "LGPD") e pela Regulação Geral de Proteção de Dados da União Europeia (aqui sob o acrônimo de "GDPR"), que, todavia, impõem limitações que podem ter impacto no seu efetivo exercício.

Todavia, não há consenso na doutrina sobre a real existência de um direito à explicação. Selbst e Powles[12] defendem a existência efetiva de tal direito, em claro contraponto a outros autores como Watcher, Mittelstadt, Floridi[13] e Russel[14]. Os que defendem a existência afirmam categoricamente que a GDPR, ao estabelecer direitos

---

[8] "A proteção de dados pode ser vista como a soma de um conjunto de direitos que configuram a cidadania do novo milênio". Cf.: RODATÀ, Stefano. A vida na sociedade da vigilância. tradução: Danilo Doneda e Luciana Cabral Doneda. Rio de Janeiro: Renovar, 2008, p. 17.

[9] O'NEIL, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. New York: Crown.

[10] PASQUALE, F. (2016). *The blackbox society: The secret algorithms that control money and information*. Cambridge: Harvard University Press.

[11] WACHTER, S.; MITTELSTADT, B.; FLORIDI, L. (2017). "Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation". International Data Privacy Law. Disponível em: https://ssrn.com/abstract=2903469. Acesso em: 02 jul. 2021.

[12] SELBST, A. D.; POWLES, J. (2017). "Meaningful information and the right to explanation". International Data Privacy Law, vol. 7, nº 4, p. 233-242. Disponível em: https://ssrn.com/abstract=3039125 Acesso em: 02 jul. 2021.

[13] Op. cit.

[14] MITTELSTADT, Brent; RUSSELL, Chris; WACHTER, Sandra. Explaining Explanations in AI. Rochester, NY: Social Science Research Network, 2018. Disponível em: https://papers.ssrn.com/. Acesso em: 02 jul. 2021.

de informação sobre a lógica de processos de decisões automatizadas[15], confere claramente o direito à explicação, e este deve ser interpretado de modo a permitir ao titular dos dados o exercício de seus direitos previstos na GDPR e no ordenamento jurídico.[16] Os que se opõem argumentam que a ausência do termo “explicação” no texto da GDPR não permitiria afirmar categoricamente a existência de tal direito na amplitude semântica do termo.[17]

Neste trabalho, a hipótese é que tal direito, no contexto da LGPD, existe, numa proporção até maior do que na GDPR, devido à forma como o princípio da transparência, obrigações de informação e o próprio conceito de dado pessoal foram adotados pela legislação geral nacional. Todavia, a sua implementação prática encontra diversos obstáculos, como os limites estabelecidos por regras de segredo de negócio e propriedade intelectual, em paralelo à complexidade cada vez maior dos algoritmos e sua opacidade quase que natural.[18] Essa dificuldade de implementação pode dificultar a garantia de exercício de outros direitos individuais e coletivos, exacerbar a assimetria de informação e por consequência a assimetria de poder, além de complexificar a demonstração de eventual responsabilidade por práticas inadequadas e abusivas no tratamento de dados pessoais.

Portanto, neste trabalho, partimos de uma preocupação em abordar e analisar tais potenciais insuficiências e dificuldades em garantir o direito à explicação, culminando em propostas capazes de saná-las. O capítulo um desta tese apresenta o problema, a hipótese e as lacunas do debate, bem como apresenta algumas definições prévias para a nossa reflexão, que são os conceitos de algoritmo e do que é decidir e o que são decisões. Este tema é polêmico, de forma que tais definições são meramente instrumentais. Optou-se pelas propostas que permitissem a compreensão do problema do direito à explicação, sem com isso apontar para uma resposta definitiva sobre os problemas éticos e filosóficos destes conceitos. Por fim, apresentamos uma breve consideração metodológica sobre o desenvolvimento da pesquisa, como forma de orientar o leitor sobre os limites e alcance das conclusões

---

[15] Os artigos 13 e 14 da GDPR garantem o direito à “informações úteis relativas à lógica subjacente”.
[16] SELBST, A. D.; POWLES, J., op. cit.
[17] WACHTER, S; MITTELSTADT, B.; FLORIDI, L. 2017, op. cit.
[18] BURRELL, J. How the machine ‘thinks’: Understanding opacity in machine learning algorithms. *Big Data & Society*. Disponível em: http://ssrn.com/abstract=2660674. Acesso em: 02 jul. 2021.

deste trabalho.

O capítulo dois apresenta o debate sobre a existência do direito à explicação no contexto europeu e as possíveis interpretações do Art. 22 do GDPR e sua relação com os demais elementos do regulamento que disciplinam as obrigações de transparência e os direitos dos titulares de dados. Dessa forma, discutiremos como esse debate passou a se apresentar no contexto brasileiro a partir da 13.709 de 2018, a Lei Geral de Proteção de Dados (LGPD) bem como de outras normas e precedentes judiciais relacionados à proteção de dados, transparência algorítmica e *accountability*. O objetivo do capítulo é apresentar o que seriam decisões automatizadas, o que seria a explicação no contexto de decisões automatizadas e as diferentes formas pelas quais se pode fundamentar o direito à explicação. A fundamentação que apresentaremos baseou-se em uma interpretação sistemática do ordenamento brasileiro, incluindo discussões sobre os direitos da personalidade, a autodeterminação informacional, a garantia de decisões adequadas e o devido processo informacional.

O capítulo dois serve como um primeiro passo na compreensão do problema, que será aprofundado no capítulo três, a partir de uma análise da explicabilidade de decisões automatizadas em diferentes contextos, quais sejam, os EUA, a Europa e o Brasil. Essa análise permite compreender de forma mais precisa os problemas relacionados à regulações específicas que buscam endereçar o problema da governança algorítmica, bem como os contornos mais precisos sobre quais os dispositivos sobre decisões automatizadas podem ser encontrados em cada um dos ordenamentos jurídicos.

A partir dessa análise dos textos legais, é possível compreender seu escopo e incidência, bem como as limitações que a aplicação desse direito pode encontrar na prática. Essas limitações serão aprofundadas no capítulo quatro. O debate em torno da transparência algorítmica apresenta três elementos como os principais desafios. O primeiro deles são as limitações legislativas em torno do segredo de negócios e do conceito de dados pessoais.

Os segredos de negócio, previstos no ordenamento Europeu como “direitos de terceiros” e na LGPD como “segredo comercial e industrial”, aparecem como

limitações ao direito à explicação. Se o direito de acesso permite ao sujeito tomar conhecimento se determinados dados pessoais encontram-se em poder de terceiros, o mesmo não ocorre com os critérios utilizados em determinado processamento de dados, visto que tais informações podem revelar estratégias comerciais.

Outro desafio legislativo relevante diz respeito à limitação do conceito de dados pessoais. A definição presente no GDPR e na LGPD apresenta o conceito de dado pessoal como relacionado a "pessoa natural identificada ou identificável", no entanto, a natureza de algumas inferências estatísticas pode permitir um vazio regulatório em torno de alguns dados que não possuam relações diretas como pessoas identificáveis, mas que por sua vez podem apresentar impactos significativos na esfera de direitos individuais ou coletivos[19].

Outro elemento que será discutido serão as limitações que chamaremos de cognitivas. Conforme divisão apresentada por Jenna Burrel,[20] é possível observar três dificuldades para a compreensão do processamento algorítmico de dados, ou, nos termos da autora, três tipos de opacidade. A primeira é a já mencionada limitação em relação ao segredo de negócio. A segunda é a natureza técnica dos sistemas algorítmicos, que dependem de conhecimentos matemáticos, sobre programação e sobre tecnologias da informação para serem plenamente compreendidos. A terceira forma é relacionada a determinadas aplicações de *machine learning* que possuem uma opacidade intrínseca ao seu modo de funcionamento, visto operarem volumes tão grandes de operações matemáticas, de formas não relacionadas a lógica utilizada no raciocínio humano. Embora possa se conhecer os outputs desses sistemas, as operações realizadas pelo algoritmo não são interpretáveis nem mesmo pelos próprios programadores. A esses dois últimos tipos de opacidade apontado por Burrel, acrescentamos mais um limite, que também possui uma natureza cognitiva, que diz respeito à realidade socioeconômica, visto que o conhecimento sobre tecnologias é desigualmente distribuído entre a população, fenômeno que é debatido como o *digital divide*.

---

[19] WACHTER, S.; MITTELSTADT, B. *A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI*. [*s. l.*], 2018. Disponível em: https://doi.org/10.31228/osf.io/mu2kf. Acesso em: 27 maio 2020.

[20] BURRELL, J. How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*. Disponível em: http://ssrn.com/abstract=2660674. Acesso em: 02 jul. 2021.

Por fim, como outro desafio ao direito à explicação, discutiremos as limitações institucionais decorrentes da capacidade de atuação das entidades supervisoras de proteção de dados em seu papel normativo e fiscalizador, tais como a Autoridade Nacional de Proteção de Dados (“ANPD”), no Brasil.

Uma vez apresentados os limites do direito à explicação, o capítulo cinco discute os instrumentos jurídicos envolvidos na garantia e implementação do direito, começando pela discussão das responsabilidades dos agentes de tratamento. Após essa discussão, apresentamos as obrigações relacionadas à transparência e as diferentes formas de exercê-la, seja de forma passiva, em resposta às demandas subjetivas, ou de forma ativa, implementando medidas para garantir informações acessíveis.

Apontado como um dos princípios mais relevantes na garantia da proteção de dados, e principalmente nos processos de decisões automatizadas, o conceito de *accountability*[21], presente na LGPD como prestação de contas e responsabilidade demonstrável, indica uma uma série de medidas necessárias para a implementação de sistemas que utilizem algoritmos para o tratamento automatizado de dados pessoais, incluindo a obrigação de prestar contas sobre o tratamento. Discutiremos em que medida as organizações devem prestar contas às entidades e aos próprios titulares de dados, principalmente na realidade de assimetrias informacionais.

Outro instrumento que discutiremos será o Relatório de Impacto a Proteção de Dados, sua importância nos contextos de desenvolvimento de decisões automatizadas, bem como outras modalidades de avaliação, alternativas ou subsidiárias, que podem contribuir no desenvolvimento de sistemas de decisões automatizadas, para garantir decisões mais justas e menos propensas a vieses e impactos sociais. Além disso, como essas avaliações podem contribuir para informar os riscos das aplicações e quais são os elementos que precisam constar em uma explicação. Na mesma linha, apresentaremos uma discussão em torno do conceito de *explainability by design*, como consequência do entendimento de que a explicabilidade

[21] ALHADEFF, J.; VAN ALSENOY, B.; DUMORTIER, J. The Accountability Principle in Data Protection Regulation: Origin, Development and Future Directions. *In*: GUAGNIN, D. *et al.* (org.). *Managing Privacy through Accountability*. London: Palgrave Macmillan UK, 2012. p. 49–82. Disponível em: https://doi.org/10.1057/9781137032225_4. Acesso em: 16 dez. 2020.

deve estar contida nas medidas necessárias desde a previsão de *privacy by design* presente no GDPR e na LGPD.

Como último instrumento para garantia do direito à explicação, discutiremos a auditoria. A partir da sua apresentação no campo da proteção de dados, discutiremos como podem ser utilizadas pelas autoridades de proteção de dados, as possíveis abordagens e limitações, bem como as possibilidades de que sejam realizadas de forma preventiva pelas próprias organizações.

Por fim, como conclusão do capítulo, apresentaremos um *framework* para aplicação e mensuração prática do direito à explicação, utilizando os instrumentos debatidos no trabalho e propondo uma possível abordagem para o devido processo informacional no contexto de decisões automatizadas.

1 PROBLEMA, HIPÓTESE, LACUNA E ORIGINALIDADE

O problema de pesquisa pode ser formulado a partir de uma simples pergunta: existe um direito à explicação no contexto da Lei Geral de Proteção de Dados do Brasil? Essa pergunta remete-nos a um debate oriundo do outro lado do Atlântico. Desde a aprovação do GDPR, o debate em torno do art. 22 do Regulamento, que previa o direito à revisão de decisões automatizadas, chamou a atenção dos estudiosos da proteção de dados do contexto europeu, pela previsão de que os agentes de tratamento deveriam fornecer aos titulares oportunidades de questionar a decisão

Embora não exista no Regulamento uma previsão expressa de um direito à explicação, alguns autores afirmam que a interpretação do art. 22, em conjunto com seus *recitals* e os artigos relacionados às obrigações de transparência, permitiam afirmar que o titular de dados possui um direito à explicação no direito da União Europeia, visto que é necessário compreender o funcionamento do sistema responsável pelo processamento automatizado de dados para efetivamente exercer o direito de desafiar tal decisão.

A redação do art. 20 da LGPD acompanhou a disposição europeia de prever um direito à revisão de decisões automatizadas. O mesmo artigo prevê que os responsáveis pela decisão forneçam informações claras sobre os critérios e procedimentos da decisão. A principal hipótese desta pesquisa, baseada na revisão de literatura apresentada, e nas atuais discussões, é que, sim, existe fundamentado na Lei Geral de Proteção de Dados do Brasil e na cláusula geral do devido processo legal, um direito à explicação no contexto de decisões automatizadas orientadas por algoritmos que se baseiam no tratamento de dados pessoais.

No Brasil, desde a aprovação da LGPD, embora a discussão tenha repercutido no debate público, poucos são os estudos que se debruçaram sobre a existência do direito à explicação no direito brasileiro de forma aprofundada. De acordo com o banco de teses e dissertações da Biblioteca Brasileira de Teses e Dissertações, o tema aparece em apenas um trabalho sobre a transparência algorítmica, mas não o aborda de forma específica, apenas como uma das questões atinentes ao princípio da transparência[22]. No contexto internacional, em pesquisa realizada no *Open Access*

[22] FLORÊNCIO, J. A. *Proteção de dados na cultura do algoritmo*. 2019. 320 f. Tese (Doutorado em

Theses and Dissertations[23] por meio das palavras chaves “right to explanation”, apenas duas teses foram encontradas, uma da Universidade de Tilburg[24], na Holanda, e uma na Universidade de Helsinki[25], na Finlândia.

A principal lacuna na compreensão do problema diz respeito à implementação prática desse direito. Ainda é incerto como, na prática, instrumentalizar tal direito, levando em consideração: (i) os limites impostos pela própria legislação, como segredo de negócio e propriedade intelectual; (ii) limitações cognitivas para a compreensão da tecnologias computacionais por não especialistas, e em alguns casos, até mesmo para especialistas da área, como é o caso de algumas aplicações de inteligência artificial; e, por fim, (iii) as limitações institucionais das entidades responsáveis pela supervisão de tais sistemas, que não dispõem de instrumental técnico-jurídico para *enforcement* desse direito.

Nesse sentido, a originalidade deste trabalho consiste na investigação desse problema de forma mais detalhada, na apresentação dos desafios para a sua implementação em face das limitações legais e técnicas impostas pela legislação e pelos algoritmos e na proposição de formas de instrumentalização desse direito.

1.1 ENUNCIAÇÃO DE PRESSUPOSTOS

1.1.1 O que são "dados pessoais"? Uma visão consequencialista e a necessidade de um novo conceito

Considerando que o presente trabalho propõe-se a analisar a existência de um direito à explicação no contexto de decisões automatizadas baseadas no tratamento de dados pessoais, cumpre-nos apresentar, preliminarmente, o que entendemos por “dados pessoais”. Para tanto, ao longo desta seção descreveremos as possíveis

Direito) — Programa de Estudos Pós-Graduados em Direito, Pontifícia Universidade Católica de São Paulo, São Paulo, 2019.

[23] Última pesquisa feita no endereço https://oatd.org/, no dia 20 de junho de 2021.

[24] GUNST, H. *The Right to Explanation and the Right to Secrecy* – Reconciling Data Protection and Trade Secret Rights in Automated Decision-making. 2017. Dissertation (Masters Thesis in Law) – Faculty of Law, University of Helsinki, Helsinki, 2017. Disponível em: http://hdl.handle.net/10138/231948. Acesso em: 20 jun. 2021.

[25] JANSSEN, J. H. N. *The right to explanation:* means for ‘white-boxing’ the black-box?: research into the ability of the 'right to explanation' about decisions based solely on automated decision-making of Articles 13(2)(f), 14(2)(g), 15(1)(h) and 22(3) of the General Data Protection Regulation, as well as of current explanation methods, to solve the legal problems arising from algorithmic decision-making. Dissertação (Masters Thesis in Law and Technology) — Universiteit van Tilburg, Tilburg, 2019. Disponível em: https://tilburguniversity.on.worldcat.org/search?queryString=scr.uvt.nl:8107234. Acesso em: 20 jun. 2021.

concepções de dados pessoais, inclusive além do conceito expressamente previsto na LGPD, evidenciando em que medida elas nos permitem alargar ou reduzir o escopo do conceito de dado pessoal e, por extensão, do próprio direito à explicação.

Inicialmente, cabe destacar que o conceito de dado pessoal adotado por determinada legislação carrega consigo importantes consequências práticas, uma vez que é ele quem define, em última instância, o escopo da tutela conferida aos titulares.[26] Tradicionalmente, as leis gerais de proteção de dados pessoais costumam oscilar entre as abordagens expansionista, caracterizada pela adoção de um vocabulário que alarga o conceito de dado pessoal e, consequentemente, o alcance da tutela conferida ao titular e a abordagem reducionista, que restringe os contornos do conceito de dado pessoal, limitando o escopo de proteção.[27]

Em linhas gerais, partindo-se de uma abordagem reducionista, consideram-se dados pessoais unicamente as informações relacionadas a uma pessoa *identificada*, devendo haver, portanto, um vínculo imediato, direto, preciso e exato entre o dado e o titular. A abordagem reducionista traz consigo, portanto, uma retração da qualificação do dado como pessoal. A abordagem expansionista, por sua vez, permite definir dado pessoal como a informação relacionada a uma pessoa natural identificada ou *identificável*. Com base nesta abordagem, é possível que o titular seja uma pessoa indeterminada e que o vínculo estabelecido entre o dado e aquele seja mediato, indireto, impreciso e inexato, permitindo um alargamento da qualificação do dado como pessoal.[28]

A Lei Geral de Proteção de Dados do Brasil (Lei nº 13.709/2018), em linha com a maior parte das leis gerais de proteção de dados pessoais ao redor do mundo e dos instrumentos internacionais vigentes[29], adota o conceito expansionista, sem rol exemplificativo, ao definir dado pessoal como a informação relacionada a pessoa natural identificada ou *identificável* (art. 5º, I).

Em contraposição, os dados anonimizados seriam a antítese do conceito de

---

[26] BIONI, B. Xeque-Mate: o tripé de proteção de dados pessoais no xadrez das iniciativas legislativas no Brasil. *GPOPAI/USP*, [*S. l.*], 2015. Disponível em: https://www.researchgate.net/publication/328266374_Xeque-Mate_o_tripe_de_protecao_de_dados_pessoais_no_xadrez_das_iniciativas_legislativas_no_Brasil. Acesso em: 27 ago. 2020. p. 17; e BIONI, B. R. *Proteção de dados pessoais*: a função e os limites do consentimento. 2. ed. Rio de Janeiro: Forense, 2021. p. 59.

[27] Idem. *Proteção de dados pessoais*: a função e os limites do consentimento. 2. ed. Rio de Janeiro: Forense, 2021.

[28] BIONI, B. R. *Proteção de dados pessoais*: a função e os limites do consentimento. 2. ed. Rio de Janeiro: Forense, 2021.

[29] Ibidem, p. 65.

dado pessoal. Em resumo, podem ser entendidos como dados anonimizados aqueles incapazes de revelar a identidade de uma pessoa após passarem por um processo de quebra do vínculo entre o dado e o seu titular (processo de anonimização), que pode empregar diferentes técnicas, variando entre supressão, generalização, randomização e pseudoanonimização.[30] Por muito tempo acreditou-se na ideia de que seria possível uma completa e irreversível anonimização, entendimento que se revelou equivocado, havendo vários estudos empíricos que o contestam, evidenciando a natureza tecnologicamente imperfeita do processo de anonimização, que pode ser revertido, por exemplo, por meio da combinação de diferentes bases de dados ou de pequenos *bits* de informação.[31] Neste sentido assinala Bioni:

> A proteção dos dados pessoais, como um novo direito da personalidade, dirige-se a todo e qualquer dado em que se denote o prolongamento de um sujeito. Dados pessoais não se limitam, portanto, a um tipo de projeção imediata, mas, também, a um referencial mediato que pode ter ingerência na esfera de uma pessoa. *Por essa lógica, qualquer dado pessoal anonimizado detém o risco inerente de se transmudar em um dado pessoal*. A Agregação de diversos "pedaços" de informação (dados) pode revelar (identificar) a imagem (sujeito) do quebra-cabeça, a qual era até então desfigurada (anônimo) – o chamado efeito mosaico.[32] (grifo nosso)

De acordo com Bioni, leis que adotam um conceito expansionista de dados pessoais e ao mesmo tempo estabelecem uma dicotomia destes com os dados anonimizados poderiam incorrer numa incoerência. A solução encontrada pela LGPD foi a adoção de um filtro para delimitar melhor o escopo do conceito expansionista (filtro da razoabilidade): um dado deixará de ser considerado anonimizado apenas se o processo de anonimização puder ser revertido mediante o emprego de esforços razoáveis[33], nos termos do art. 12:

> Art. 12. Os dados anonimizados não serão considerados dados pessoais para os fins desta Lei, salvo quando o processo de anonimização ao qual foram submetidos for revertido, utilizando exclusivamente meios próprios, ou quando, com esforços razoáveis, puder ser revertido.
> § 1º A determinação do que seja razoável deve levar em consideração fatores objetivos, tais como custo e tempo necessários para reverter o processo de anonimização, de acordo com as tecnologias disponíveis,

[30] BIONI, B. R. *Proteção de dados pessoais*: a função e os limites do consentimento. 2. ed. Rio de Janeiro: Forense, 2021. p. 61-62.
[31] Ibidem, p. 63-64.
[32] Ibidem, p. 65.
[33] Ibidem, p. 66.

> e a utilização exclusiva de meios próprios.
> § 2º Poderão ser igualmente considerados como dados pessoais, para os fins desta Lei, aqueles utilizados para formação do perfil comportamental de determinada pessoa natural, se identificada.

A calibração desse filtro de razoabilidade, de acordo com o autor, dá-se a partir do estabelecimento de critérios objetivos e subjetivos como fatores de uma análise de risco. Neste sentido, aponta-se que a adoção do termo "razoável" revela a intenção do regulador em buscar uma norma tecnologicamente neutra, isto é, não vinculada a uma tecnologia específica e que não se tornasse ultrapassada com o seu desenvolvimento. Em contrapartida, preferiu-se adotar um conceito aberto, associando a ele algumas balizas para reduzir a discricionariedade em sua interpretação. Sendo assim, adotaram-se como critérios objetivos a análise do estado da arte da tecnologia, avaliada a partir dos fatores *custo* e *tempo*, e ainda um segundo eixo de análise, de caráter subjetivo. Neste eixo, "deve-se levar em conta quem é o agente de tratamento de dados e se ele dispõe de "meios próprios" para reverter o processo de anonimização", considerando-se: (i) o fluxo de dados dentro de uma organização e (ii) o fluxo de dados para fora da organização.[34] Nas palavras do autor:

> Em síntese, o legislador brasileiro adotou uma estratégia normativa alinhada à premissa de que os dados anonimizados seriam sempre passíveis de reversão. Os dois eixos de análise acima descritos – objetivo e subjetivo – compõem uma matriz de risco em torno de possíveis engenharias reversas de um processo de anonimização. A resiliência de tal processo é o que determinará se haverá algum tipo de intersecção entre dados anonimizados e dados pessoais, cujos elementos de análise são de ordem objetiva (razoabilidade) e subjetiva (meios próprios).[35]

Partindo-se da noção de que todo dado anonimizado é um dado pessoal em potencial, com base na ideia do filtro de *razoabilidade*, observamos que a fronteira entre o conceito de dado pessoal e de dado anonimizado não é tão simples nem tão clara. Essa zona cinzenta entre um domínio e outro é ainda mais aprofundada na LGPD. A redação do art. 12 da LGPD, em especial seu § 2º, ao estabelecer que "Poderão ser igualmente considerados como dados pessoais, para os fins desta Lei, aqueles utilizados para formação do perfil comportamental de determinada pessoa

[34] BIONI, B. R. *Proteção de dados pessoais*: a função e os limites do consentimento. 2. ed. Rio de Janeiro: Forense, 2021. p. 68-71.
[35] Ibidem, p. 71-72.

natural, se identificada”[36], dá origem a uma terceira abordagem de definição do conceito de dado pessoal, cunhada como abordagem consequencialista, que representa uma verdadeira zona cinzenta entre as noções de dado pessoal e dado anonimizado.

A abordagem consequencialista coloca em perspectiva não a possibilidade de estabelecimento de um vínculo entre determinado dado e uma pessoa natural identificada ou identificável, mas a própria possibilidade de um determinado dado ou tratamento afetar um indivíduo ou grupo. Sendo assim, partindo-se da perspectiva consequencialista, torna-se irrelevante saber se é possível estabelecer um vínculo, seja ele mediato ou imediato, direto ou indireto, entre dado e pessoa natural identificada ou identificável. Se o tratamento em questão for relevante e impactar de forma significativa o livre desenvolvimento da personalidade de um grupo ou indivíduo, os dados empregados, ainda que anonimizados, poderão ser considerados dados pessoais para os fins da lei, nos termos do art. 12, § 2º. Em outras palavras, na abordagem consequencialista o foco está muito mais na consequência e no impacto que o tratamento de um dado pode ter na vida de um determinado sujeito, população ou grupo de pessoas (análise de *causa e efeito*) do que no estabelecimento de um vínculo entre dado e pessoa natural identificada ou identificável. Nas palavras de Bioni:

> Se a premissa da causa regulatória da proteção de dados pessoais é tutelar o cidadão, que é cada vez mais exposto a tais tipos de práticas que afetam a sua vida, então, uma compartimentalização “dura” entre dados pessoais e dados anonimizados deixaria de fazer sentido. Em especial, quando está em questão a formação de perfis comportamentais que tem por objetivo precípuo influenciar de alguma forma a vida de uma pessoa, que está atrás de um dispositivo e pouco importa ser ela identificável ou não. *Abre-se espaço, assim, para uma escolha normativa consequencialista. Não se normatiza apenas pela lente da conceituação mutuamente excludente entre dados pessoais e dados anônimos, mas, também, por meio da relação de causa e efeito que a mera atividade de tratamento de dados pode exercer sobre um indivíduo. Não se deve perder de vista, portanto, que mesmo o tratamento de dados anonimizados pode repercutir na esfera do livre desenvolvimento da personalidade das pessoas*. Algoritmos que mineram dados anonimizados podem esconder práticas discriminatórias em prejuízo de uma coletividade e de pessoas

[36] BRASIL. *Lei nº 13.709, de 14 de agosto de 2018*. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF: Presidência da República, 14 ago. 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015- 2018/2018/lei/L13709.htm. Acesso em: 07. nov. 2020.

> singulares.[37] (grifo nosso)

Conforme assinala Bioni, a abordagem consequencialista plasmada na LGPD traz consigo relevantes consequências de ordem pragmática, permitindo alargar significativamente o escopo da proteção de dados pessoais, tornando possível uma alocação dogmática do direito à proteção de dados pessoais como um novo direito da personalidade oponível não só a indivíduos, mas também a grupos. De acordo com o autor, pouco importa se um tratamento utiliza uma informação isolada ou agregada e que não esteja associada direta ou indiretamente a uma pessoa identificada ou identificável. Esse tratamento será relevante sempre que impactar de forma significativa o livre desenvolvimento da personalidade de um indivíduo.[38] Em outras palavras, se o uso de dados, que inclusive podem ser não pessoais, impactar direitos e liberdades fundamentais de um titular ou de um grupo. A adoção de um conceito calcado numa abordagem consequencialista de dados pessoais traz consigo a possibilidade de uma maior tutelabilidade de relações baseadas no tratamento de dados, sejam eles pessoais ou não, o que se mostra especialmente relevante no contexto de decisões automatizadas, que nem sempre se valem de dados pessoais na acepção tradicional do termo, mas também de outros tipos de dados que, ainda que não pessoais, acabam impactando grupos e/ou indivíduos:

> Daí a importância da alocação da proteção de dados pessoais como um novo direito da personalidade. Com isso, permite-se um alcance normativo maior, que é capaz de abraçar toda e qualquer atividade de processamento de dados (ainda que não pessoal), mas que impacta a vida de um indivíduo.[39]

Conforme argumenta Bioni, essa racionalidade de viés consequencialista é adotada no art. 12, § 2º, da LGPD, que coloca o foco nas consequências que um determinado tratamento pode acarretar para um sujeito. De acordo com o autor, por vezes, decisões automatizadas valem-se não de perfis individuais, mas de *perfis de grupos (grouping),* pelo que se deve entender que as expressões *determinada pessoa* e *identificada*, presentes no art. 12, § 2º, devem ser interpretadas sob essa ótica de que determinado tratamento pode vir a impactar um indivíduo, e não como expressões

---

[37] BIONI, B. R. *Proteção de dados pessoais*: a função e os limites do consentimento. 2. ed. Rio de Janeiro: Forense, 2021. p. 76-77.
[38] Ibidem, p. 77.
[39] Ibidem, p. 78.

relacionadas a uma pessoa identificada em específico.[40] De acordo com o autor, essa interpretação sistemática do art. 12, § 2º, está em linha com o conceito expansionista e com um dos objetivos e fundamentos da lei, que é o livre desenvolvimento da personalidade (arts. 1º e 2º, VII). Ainda segundo o autor, essa interpretação sistemática seria reforçada quando analisada a redação original do art. 12 no Projeto de Lei de Proteção de Dados Pessoais de autoria do Poder Executivo (PLPDP/EXE) (versão de out. 2015), que estipulava: "Poderão ser igualmente considerados como dados pessoais para os fins desta Lei os dados utilizados para a formação do perfil comportamental de determinadas modalidades uma determinada pessoa natural, *ainda que não identificada.*"[41] (grifo nosso). O autor ainda acrescenta que essa interpretação é sustentada por um quadro de elementos normativos, conformado, como vimos, pelo art. 12 do PLPDP/EXE, e pelos arts. 12, § 2º, art. 20, art. 5º, I, e art. 2º, VIII, da LGPD.[42]

No presente trabalho, alinhamo-nos a uma abordagem consequencialista de dados pessoais e propugnamos pela necessidade de uma regulação baseada no impacto às pessoas e não apenas na proteção dos dados pessoais *per se*. Apesar de as recentes gerações de leis de proteção de dados terem avançado neste aspecto[43], elas continuam fadadas a falhar no cumprimento de seu objetivo básico: proteger a privacidade e outros direitos fundamentais como direitos humanos. Isso porque elas continuam a proteger os dados em si, e não os impactos que o uso indevido de dados pode ter nas pessoas, o que se torna cada vez mais relevante no contexto da proliferação de sistemas que não precisam necessariamente processar dados pessoais para ter impacto nas pessoas, a exemplo de muitos modelos de negócio e tecnologias de publicidade digital.[44] No momento, as grandes empresas estão investindo em tecnologias — como privacidade diferencial, dados sintéticos e aprendizado de máquina — que podem processar grandes quantidades de dados não pessoais e, ainda assim, afetar as pessoas. A próxima geração de leis de proteção de

---

[40] BIONI, B. R. *Proteção de dados pessoais*: a função e os limites do consentimento. 2. ed. Rio de Janeiro: Forense, 2021.

[41] INTERNETLAB. O que está em jogo no debate sobre dados pessoais no Brasil? Relatório final sobre o debate público promovido pelo Ministério da Justiça sobre o Anteprojeto de Lei de Proteção de Dados Pessoais. São Paulo: InternetLab, 2016. p. 158.

[42] Ibidem, p. 79.

[43] Sobre a análise geracional das leis de proteção de dados pessoais, cf. nota de rodapé nº 69.

[44] RAVICHANDRAN, D.; VASSILVITSKII, S. Evaluation of Cohort Algorithms for the FLoC API. [*S. l.*], [*s. n.*], 21 out. 2020. Disponível em: https://github.com/google/ads-privacy/blob/master/proposals/FLoC/FLOC-Whitepaper-Google.pdf.

dados precisa se concentrar no impacto que o processamento de dados, pessoais ou não pessoais, pode ter sobre os direitos fundamentais das pessoas.[45] A LGPD, como vimos, possui alguns dispositivos que podem ser interpretados neste sentido, sendo possível afirmar que ela representaria uma "geração 4.5" das leis de proteção de dados pessoais, ao representar uma fronteira entre o paradigma consolidado da proteção de dados calcada na proteção do dado em si e um novo paradigma no qual a regulação estaria baseada no impacto à pessoa humana. Para tanto, é essencial adotar como ponto de partida uma abordagem consequencialista de dados pessoais, que permite um alargamento da proteção conferida aos indivíduos. Nas palavras de Bioni:

> Com isso, facilita-se, dentre outras coisas, a percepção de que o tratamento de dados – sejam eles anônimos ou pessoais – que submeta uma coletividade ou uma pessoa a processos de decisões automatizadas deve estar dentro do escopo normativo da proteção dos dados pessoais. Essa é uma chave de leitura essencial para a compreensão da matéria na cultura jurídico-legal brasileira e dos desafios regulatórios de uma sociedade e uma economia cada vez mais movidas por dados.[46]

Nesta tese, portanto, partimos desse paradigma conceitual mais amplo, que dá origem à expectativa de reconhecimento de um direito à explicação de escopo mais alargado do que aquele possível no cenário europeu, que englobaria não apenas decisões automatizadas baseadas no tratamento de dados pessoais, mas toda e qualquer decisão automatizada que possa vir a impactar a esfera de direitos de um titular ou um grupo, ainda que não baseada no tratamento de dados estritamente pessoais.

1.1.2 O que são "algoritmos"? Muito além da receita de bolo

Segundo Danilo Doneda e Virgílio Almeida[47], o termo algoritmo designa simplesmente "[...] um conjunto de instruções para a realização de uma tarefa, produzindo *output* a partir de *input*." Consoante com essa noção está a definição

---

[45] TENE, Omer. Privacy: The new generations. *International Data Privacy Law*, v. 1, n. 1, p. 15-27, fev. 2011. Disponível em: https://academic.oup.com/idpl/article/1/1/15/759641. Acesso em; 21 jun. 2021.
[46] BIONI, B. R. *Proteção de dados pessoais*: a função e os limites do consentimento. 2. ed. Rio de Janeiro: Forense, 2021. p. 81.
[47] DONEDA, D.; ALMEIDA, V. What Is Algorithm Governance? *IEEE Internet Computing*, [*S. l.*], v. 20, p. 60–63, 2016. Disponível em: https://doi.org/10.1109/MIC.2016.79. Acesso em: 21 jun. 2021.

também dada por Tarteton Gillespie[48], para quem os algoritmos são um truque, são um procedimento utilizado para tornar uma tarefa operacionalizada.

Vê-se como o termo pode abarcar as funções mais básicas de um computador, como a função de calculadora, onde o usuário insere uma série de números, requisita um conjunto de operações e recebe um resultado. De fato, essa e todas as funções de um computador são operacionalizadas por meio de algoritmos que, como dito, são a peça básica de qualquer programa de computador.

No entanto, o termo algoritmo é uma palavra que entrou na moda e está presente diariamente nas mídias e nas conversas cotidianas. Seu sentido extrapolou o termo técnico, de forma que a palavra assumiu diversos sentidos. Com o crescente uso de tecnologias digitais, a penetração de processamentos de dados na realidade social é quase absoluta. Atualmente, é quase impossível ter uma interação humana que não seja de alguma forma mediada, auxiliada ou acompanhada de serviços digitais. E esses, por sua vez, são operados por algoritmos.

Tartenton Gillespie levanta alguns sentidos relevantes pelos quais a palavra é utilizada. Um deles é chamado de “Talismã”, que corresponderia a todo o campo simbólico existente em torno da palavra algoritmo, que evoca noções de objetividade e cientificismo. O termo é utilizado, por exemplo, como uma proposta de *marketing* em torno de serviços digitais mais eficientes.

O termo algoritmo ainda funciona como uma abreviação para a combinação de diferentes sistemas, *softwares*, *hardwares*, coleta e processamento de dados, modelos matemáticos e toda uma cadeia de processos digitais ou não. Ao se discutir “o algoritmo do Facebook”, na verdade não está se discutindo apenas o conjunto de operações matemáticas que operam o serviço da rede social, mas o que há por trás de todo um conjunto de decisões da empresa Facebook que esculpiram aquele serviço e que orientam tudo que o serviço faz. Ou seja, ao se referir ao algoritmo, estas utilizações comuns do termo correm o risco de perder de vista a dimensão consciente, deliberada e humana daqueles que conceberam o serviço e que o operam.

O problema deste trabalho engloba as duas dimensões, por considerar as decisões automatizadas como um momento específico de um processo sociotécnico

[48] GILLESPIE, T. Algorithm*. In:* PETERS, B. *Digital Keywords*. Princeton: Princeton University Press, 2016. p. 18-30.

complexo de relações jurídicas e sociais. No entanto, ao utilizarmos o termo algoritmo, adotaremos o sentido da primeira definição, de um conjunto de instruções logicamente ordenadas para atingir um determinado fim desejado pelo programador. A literatura costuma utilizar outros termos como sinônimos, como modelos ou sistemas.

Para tratar dos aspectos sistêmicos do processamento de dados pessoais, utilizaremos termos que denotem essa realidade. Quando falamos em “discriminação algorítmica”, por exemplo, estaremos nos referindo não a discriminação específica de um algoritmo, mas do conjunto de processamentos de dados, realizados por meio de linguagem informacional, que resultam em práticas discriminatórias e que inclui não apenas os *outputs* de um sistema, mas os fatores técnicos e sociais que tem como efeito essa discriminação.

Feita essa apresentação, podemos fornecer uma descrição mais palpável do que são algoritmos. A autora Hannah Fry[49] descreve em maiores detalhes as diferentes aplicações nas quais esse conjunto de instruções podem ser utilizadas. Em seu livro, define o algoritmo como um procedimento passo-a-passo, ordenado para solucionar um problema. Ela destaca que no sentido mais amplo, qualquer série de instruções poderia ser considerada um algoritmo. Isso inclui qualquer receita de bolo, um passo-a-passo de um tutorial do YouTube, um manual para montar um Lego ou mesmo as direções oferecidas por um transeunte, ao buscar um endereço na rua.

A autora aponta quatro atividades muito comuns que se apresentam como os principais modelos de algoritmos, apontados por Nicholas Diakopoulos[50]. Essa classificação será muito útil ao desenvolvimento deste trabalho. Em primeiro lugar, há algoritmos de priorização, que oferecem uma lista ordenada. Exemplos óbvios são o motor de buscas do Google ou a ferramenta de ordenação do Word. A ordenação obedece a critérios previamente estabelecidos, como ordem alfabética, ordem crescente de valores, entre outros, que podem ser úteis para uma atividade.

Em segundo, temos algoritmos de categorização. Esses encontram a categoria adequada para determinado item, dentro de uma tipologia. Por exemplo, podem classificar Maria segundo uma série de critérios: sexo, idade, cidade de residência. Esses são os algoritmos responsáveis por fazer a marcação automática de fotos, por

---

[49] FRY, H. *Hello World*: How to be Human in the Age of the Machine. London New York Toronto Sidney Auckland: Doubleday, 2018.

[50] DIAKOPOULOS, N. *Algorithmic accountability reporting:* on the investigation of black boxes. [*S. l.*]: Tow Center for Digital Journalism, 2014. Disponível em: https://academiccommons.columbia.edu/doi/10.7916/D8ZK5TW2. Acesso em: 21 jun. 2021.

exemplo, ou que operam os serviços de rastreamento que nos oferecem publicidade direcionada. Esse tipo de atividade, por exemplo, pode ser utilizado em veículos autônomos, para reconhecer objetos no ambiente.

Em terceiro lugar há os algoritmos de associação, que encontram relações entre dois ou mais inputs. São esses algoritmos que sugerem, por exemplo, produtos na Amazon, músicas no Spotify ou vídeos no Youtube. Esse tipo de operação é útil por encontrar padrões e relações entre diferentes fatores, as quais muitas vezes não seriam facilmente descobertas por seres humanos.

Em quarto, Fry aponta os algoritmos de filtragem, que ajudam a remover o ruído e extrair as informações importantes para determinado fim. Por exemplo, o *news feed* das plataformas de rede social que tem de filtrar o conteúdo que o usuário verá ou os serviços de comando de voz que discernem entre a voz do usuário e o barulho de fundo.

O exercício mental de imaginar combinações das atividades descritas acima permite imaginar uma série de aplicações muito úteis. A partir de um problema, é possível pensar em diferentes caminhos, tendo em vista aquelas operações, para se chegar a uma solução. A realização dessas atividades ocorre pela combinação de diferentes espécies de algoritmos, que realizam as funções mais variadas, como partes de um processo mais amplo.

As aplicações das quais trataremos neste trabalho consistem em uma combinação de diferentes algoritmos, muitas vezes desenvolvidos por diferentes atores, a partir de determinações técnicas ou comerciais, que culminam em uma aplicação específica. Um pode privilegiar celeridade, dando margem para menor precisão, enquanto outro opera de forma mais lenta e mais acurada. Dessa forma, cumpre destacar como esses processos computacionais de priorização, classificação, associação e filtragem privilegiam certos aspectos da realidade sobre outros.

A solução de um problema privilegia certo cálculo pragmático de utilidade e eficiência, de forma que algoritmos são métodos que o tempo todo fazem escolhas de custo e benefício. A escolha de um algoritmo passa, portanto, pelo resultado almejado e a finalidade que ele cumpre, assim como por variáveis contextuais, como recursos e tempo disponíveis. Quando algoritmos lidam com pessoas, contudo, a dúvida que paira no ar é sobre quais aspectos da vida individual os algoritmos estão privilegiando, ou quais aspectos não estão sequer sendo analisados e é justamente por esse motivo que a discussão da transparência dos algoritmos e a uma efetiva explicação de como

eles funcionam torna-se tão relevante.

1.1.3 O que é “decidir”? Afinal, quem decide, a máquina ou o homem?

O conceito de decisão é um dos temas mais complexos da história do pensamento. Na filosofia ocidental, reflexões sobre o problema remontam a Grécia Antiga. Depois de alguns séculos, podemos afirmar com segurança que os problemas apontados estão longe de encontrar solução. No século XX algumas teorias sobre a decisão encontraram grande repercussão a partir de estudos da economia, da ciência política e da psicologia[51]. No campo do direito há uma discussão relevante sobre a teoria da decisão judicial[52], com implicações importantes para a discussão da possibilidade de automação da atividade jurisdicional.

Discutir o conceito de decisão envolve discutir a relação entre liberdade, de um lado, como possibilidade de livre arbítrio, e de outro a necessidade de determinação dos eventos futuros pelos eventos passados, portanto, sem espaço para escolhas[53]. Essa discussão reflete no debate sobre os algoritmos e a inteligência artificial. Se a capacidade computacional permitiu que os computadores realizassem tarefas executadas por seres humanos de forma automática, há que se perguntar se os algoritmos podem decidir.

No dicionário Michaelis, encontramos alguns sentidos para o verbo decidir: apresentar um julgamento definitivo sobre algo, resolver, determinar, levar algo a uma solução, dar preferência a uma opção dentre outras possíveis[54]. Como apresentamos

---

[51] HANSSON, S. O. Decision Theory: An Overview. *In*: LOVRIC, M. (Org.). *International Encyclopedia of Statistical Science*. Berlin, Heidelberg: Springer, 2011. p. 349–355. Disponível em: https://doi.org/10.1007/978-3-642-04898-2_22. Acesso em: 14 jun. 2021; e PETERSON, M. Decision Theory: An Introduction. *In*: LOVRIC, M. (Org.). *International Encyclopedia of Statistical Science*. Berlin, Heidelberg: Springer, 2011. p. 346–349. Disponível em: https://doi.org/10.1007/978-3-642-04898-2_23. Acesso em: 14 jun. 2021.

[52] NOJIRI, Sergio. Decisão judicial. In: CAMPILONGO, C. F.; GONZAGA, A. de A.; FREIRE, A. L. (Coords.). *Enciclopédia jurídica da PUC-SP*. Tomo: Teoria Geral e Filosofia do Direito. 1. ed. São Paulo: Pontifícia Universidade Católica de São Paulo, 2017. Disponível em: https://enciclopediajuridica.pucsp.br/verbete/57/edicao-1/decisao-judicial. Acesso em: 21 jun. 2021.

[53] AZEVEDO, I. T. R.; SILVA, T. A. da. Reflexões sobre tomada de decisão e livre arbítrio sob a ótica da neurociência e seus efeitos no sistema punitivo. *LINKSCIENCEPLACE* - Interdisciplinary Scientific Journal, [*S. l.*], v. 1, n. 1, 2014. Disponível em: http://revista.srvroot.com/linkscienceplace/index.php/linkscienceplace/article/view/16. Acesso em: 14 jun. 2021; e BURNS, K.; BECHARA, A. Decision making and free will: a neuroscience perspective. *Behavioral Sciences & the Law*, [*S. l.*], v. 25, n. 2, p. 263–280, 2007. Disponível em: https://doi.org/10.1002/bsl.751. Acesso em: 28 jun. 2021.

[54] DECIDIR. *In*: MICHAELIS. São Paulo: Editora Melhoramentos, 2021. Disponível em: https://michaelis.uol.com.br/busca?r=0&f=0&t=0&palavra=decidir. Acesso em: 10 jun. 2021.

acima, algoritmos são sequências de instruções que visam a um objetivo. A programação dos algoritmos tem como consequência que o seu funcionamento deva ter por fim um resultado que pode consistir numa escolha entre várias possibilidades[55].

Podemos compreender o problema de duas formas. A primeira delas consiste na afirmação de que a escolha coube algoritmo, que considerou uma série de elementos e apresentou um resultado, ou seja, uma decisão. A segunda consiste na afirmação de que o resultado foi derivado de um processo previamente concebido, isto é, que a escolha dos códigos possíveis foi realizada por um programador, e que o algoritmo apenas seguiu as instruções e, portanto, a decisão coube ao programador.

Dessa oposição podemos deduzir ao menos duas dimensões de uma decisão. A primeira delas é a dimensão objetiva, seu conteúdo, que foi a decisão tomada pela máquina. A outra delas é mais subjetiva, e diz respeito ao processo intencional que levou a decisão tomada. A depender do conceito de decisão que adotemos, a resposta à questão da possibilidade de uma máquina decidir pode ser diferente.

Um trabalho clássico de James Moor, filósofo da ética da computação, apresenta a questão de forma muito pertinente[56]. O autor sugere que é possível dividir o conceito de decisão em dois sentidos diferentes, um mais estrito, outro mais amplo. No sentido estrito, decidir é tomar determinado caminho dentre outras possibilidades. No sentido amplo, o conceito de decisão engloba todo o processo que levou a uma decisão, que inclui a análise das possibilidades de ação, a consideração das vantagens e desvantagens e as possíveis consequências.

Diante disso, o autor questiona o que um algoritmo poderia fazer. Computadores podem certamente realizar a primeira tarefa, como escolher entre virar à direita ou à esquerda a partir de critérios definidos. Quanto ao segundo, algumas objeções são comumente levantadas sobre a possibilidade de que uma máquina possa realizar atividades outras que a resolução de problemas aritméticos ou sequências lógicas previamente delimitados. No entanto, essa visão sobre a computação não é exata. Se é verdade que os computadores apenas seguem determinado código, não é correto afirmar que todos esses códigos sejam pré-definidos.

---

[55] GILLESPIE, T. Algorithm*. In:* PETERS, B. *Digital Keywords*. Princeton: Princeton University Press, 2016. p. 18-30; MITTELSTADT, B. D. *et al.* The ethics of algorithms: Mapping the debate. *Big Data & Society*, [*S. l.*], v. 3, n. 2, 2016. Disponível em: https://doi.org/10.1177/2053951716679679. Acesso em: 20 maio 2020. p. 205395171667967.

[56] MOOR, J. Are There Decisions Computers Should Never Make. *Nature and System*, [*S. l.*], v. 1, 1985.

Uma das grandes evoluções da computação nos últimos anos diz respeito justamente à capacidade de que algoritmos produzam novos códigos e que algoritmos produzam novos algoritmos.[57] A história dos algoritmos programados para jogar xadrez pode ilustrar bem essa questão. As primeiras tentativas das máquinas jogadoras de xadrez buscavam elencar todas as possibilidades de jogadas e as probabilidades correspondentes de sucesso de cada uma delas. Nesse sentido, a avaliação dos fatores relevantes, as consequências desejadas e esperadas, as vantagens e desvantagens dessa decisão foram previamente consideradas pelo programador. Esses modelos de computador não tiveram sucesso em vencer os principais mestres humanos, mesmo que conseguissem realizar cálculos mais rápidos e precisos que eles.

Contudo, desenvolveram-se aplicações que mantêm características diferentes. Modelos onde o computador, em vez de seguir um código pré-definido, atualiza seu código com base em outros exemplos. A programação inicial não diz o que o computador deve fazer em cada caso. O programa diz apenas “como” o computador deve agir frente ao problema. A partir disso, o computador toma determinado caminho baseado no conjunto de eventos passados, que possuem uma contingência histórica, e dos estímulos externos recebidos. Quanto mais o tempo passa e mais partidas essa máquina realiza, melhor se tornam os seus movimentos no tabuleiro.

Nestas aplicações, o programador não insere previamente todos os critérios utilizados para chegar ao resultado, oferece alguns objetivos, como “cercar o rei” e operações pelas quais o programa poderá descobrir quais são os fatores mais relevantes. É um modelo desse tipo o utilizado para a criação do Deep Blue, software que conseguiu vencer o maior jogador de xadrez da época e tornou-se um marco na discussão da inteligência artificial. É esse o tipo de aplicação chamada de aprendizado de máquina (*machine learning*).

Nessa última modalidade de aplicação, Moor argumenta que é possível observarmos o segundo tipo de significado do termo decisão, visto que um sistema computacional não apenas escolheu um caminho, mas determinou os critérios relevantes, ponderou as alternativas e fez uma escolha do caminho. Nesses modelos

---

[57] STAN, F. History, motivations, and core themes. In: FRANKISH, K.; RAMSEY, W. M. (org.). The Cambridge Handbook of Artificial Intelligence. Cambridge: Cambridge University Press, 2014. Disponível em: https://doi.org/10.1017/CBO9781139046855.007. Acesso em: 24 ago. 2020; e BUCHANAN, B. G. A. (Very) Brief History of Artificial Intelligence. *AI Magazine*, [*S. l.*], v. 26, n. 4, 2005. Disponível em: https://doi.org/10.1609/aimag.v26i4.1848. Acesso em: 21 jun. 2021. p. 53.

de aplicação inclusive os critérios utilizados para a decisão são produzidos pelo algoritmo em reação aos estímulos oferecidos por informações externas, ou seja, dados de eventos anteriores.

O autor ainda rebate outra ressalva comum ao argumento de que máquinas podem decidir, visto que uma decisão precisa ser consciente. Nesse sentido, ele traz exemplos de como as pessoas tomam decisões sem reconhecer os motivos no dia a dia, como a escolha da cor da camisa, ou decidem sem nem mesmo tomar consciência de sua decisão.[58] Por fim, o autor apresenta uma distinção entre a habilidade de tomar decisões e o poder de tomar decisões, ou seja, a capacidade e a legitimidade para que essa decisão seja válida e conclui que para algumas decisões não seria prudente dar aos computadores a legitimidade para tomá-las.

A ideia de que os computadores podem decidir ainda é rebatida sob vários aspectos, principalmente sobre a dimensão moral e ética de algumas decisões. Não há muita oposição ao fato de que um algoritmo pode decidir sobre quais músicas irão para uma *playlist* com base no histórico do usuário. No entanto, ainda causa preocupação a possibilidade de que um algoritmo decida se uma pessoa é culpada ou inocente, visto que o algoritmo pode executar uma rotina de códigos e cálculos de forma rápida, mas ainda não é capaz de fazer um juízo sobre o conceito de justiça.[59]

De maneira geral, com as devidas ressalvas apontadas acima, podemos considerar que o conceito de decisão mais adequado à discussão desta tese seja aquele que se aproxima do segundo sentido apontado por Moor. Como veremos na seção 2.1.1, ao tratarmos do conceito legal de decisões automatizadas que suscitam a discussão sobre o direito à explicação, o conceito de decisão de que trata o artigo diz respeito ao resultado específico de uma aplicação e, ao mesmo tempo, os procedimentos, critérios, valores e elementos considerados para se chegar àquele resultado. Nesse sentido, podemos conceber uma decisão como a definição de um resultado possível dentro de um universo de possibilidades, incluindo o processo de

[58] O trabalho do Nobel de Economia, Daniel Kahneman, “Think Fast, Think Slow”, lançado em 2013, fornece uma perspectiva sobre os diferentes processos cognitivos, incluindo tomada de decisões e sua relação com a consciência e a intuição. Uma das grandes descobertas do autor diz respeito à descrição do funcionamento de dois módulos de pensamento distintos, um mais emocional e intuitivo, outro lógico e deliberativo.

[59] Essa questão encontra correspondência no campo das teorias da decisão jurídica. É possível identificar de um lado abordagens racionalistas, segundo as quais os aplicadores do direito apenas seguem as regras presentes na norma, que podem ser compreendidas como pressupostos de sentenças lógicas cujo sentido está contido na própria lei. Por outra lado, há abordagens hermenêuticas segundo as quais o sentido da norma precisa ser encontrado por meio da interpretação.

avaliação dos diferentes critérios e elementos objetivos e subjetivos que levaram a tal conclusão.

1.2 RECORTE METODOLÓGICO

A pergunta que orienta esse trabalho diz respeito à realidade jurídica brasileira. No entanto, ela surge no contexto europeu. Pela clara ligação dos modelos de regulação de proteção de dados adotados pelo GDPR e pela LGPD, é esperado que as discussões do velho mundo encontrem reverberações em solo tupiniquim. No entanto, a transposição de debates e interpretações exige um imenso cuidado. Nesse sentido, a importação de conceitos não ocorre de forma simples e automática de maneira que a construção da interpretação jurídica não se realiza com simplesmente um “copia e cola” de conceitos de outras jurisdições.

Isso posto, cabe realizar algumas ponderações sobre as escolhas metodológicas, tendo em vista a natureza interdisciplinar deste trabalho. A tese trata sobre a existência de um direito, dessa forma, as escolhas sobre os temas abordados se limitaram ao objetivo de esclarecer o problema do direito à explicação à luz da regulação e dos princípios da proteção de dados no contexto brasileiro. Além dos princípios, discute-se a realidade dos sistemas sociotécnicos por trás das decisões automatizadas e outras regulações que podem impactar a existência desse direito.

Este trabalho, portanto, embora realize uma importação de uma pergunta originada em um debate do continente europeu, não pode importar as respostas. Nossa resposta deve ser composta a partir do direito brasileiro, reconhecendo a influência europeia para a construção da nossa lei geral para a proteção de dados em relação com nossas instituições e tradições, e também da influência estadunidense em vários dos conceitos necessários para discutir o contexto que leve à pergunta.

Apesar das similaridades, a LGPD e o GDPR são dois regulamentos diferentes que se inserem em sistemas jurídicos distintos. Para o caso brasileiro, além dos temas comuns ao debate europeu, é preciso observar as regulações precedentes do país que trataram do tema da proteção de dados e de obrigações de transparência por tratamentos automatizados, como são o Código de Defesa do Consumidor e a Lei do Cadastro Positivo, e mesmo algumas provisões constitucionais como a previsão do *Habeas Data*.

Portanto, apesar da utilização de diferentes ordenamentos e documentos de

outras jurisdições, esta tese não é um trabalho de direito comparado. As comparações aqui estabelecidas dizem respeito às diferentes interpretações de uma previsão específica, relativa ao devido processo informacional, ao direito de revisão de decisões automatizadas e do direito que tal previsão fornece aos titulares de obterem informações sobre essa decisão. Utilizou-se também da literatura internacional para tratar de temas comuns à disciplina da privacidade e proteção de dados, sem, contudo, fazer referência e comparação a sistemas jurídicos específicos.

Desta forma, este trabalho não está comparando sistemas jurídicos, apenas analisando um fenômeno específico, utilizando-se de experiências de outros países para embasar as interpretações do sistema brasileiro. Nesse sentido, a análise da doutrina e dos precedentes são compreendidos como possibilidades de interpretação e não como respostas aos nossos problemas.

Isso posto, cabe destacar que a resposta da pergunta sobre a existência do direito à explicação envolve questões interpretativas que incluem a legislação, princípios do direito e doutrina sobre a proteção de dados. Nesse sentido, é útil para a compreensão do problema desta tese recorrer a uma visita às discussões anteriores a essas legislações específicas, seja no campo da proteção de dados, do direito constitucional, direitos fundamentais e também no campo da ciência da computação.

Os estudos sobre transparência algorítmica remontam a períodos anteriores à aprovação do GDPR e se referiam a diferentes contextos jurídicos. É possível identificar outras abordagens para a proteção de dados. De forma paralela, nos EUA, autores se debruçaram sobre o problema da transparência algorítmica a partir da ótica do devido processo legal.

O debate jurídico apresentava desde o início uma interlocução com estudos de sistemas informacionais relacionados ao campo conhecido como *Fairness, Accountability and Transparency.* Um dos grandes temas debatidos pelo campo científico da ciência da computação e do direito relacionava-se à questão dos riscos associados à opacidade de sistemas informacionais e maneiras de superá-la.

No entanto, desde a aprovação do GDPR e a previsão da decisão de decisões automatizadas presentes no art. 22, autores relevantes no cenário europeu apontavam para a questão da existência ou não do direito à explicação no regulamento. Desde então, a questão do direito à explicação e da explicabilidade de sistemas computacionais tornaram-se centrais nos campos relacionados ao direito e à tecnologia. Diversos trabalhos na área da ciência da computação passaram a

discutir a explicabilidade de modelos de inteligência artificial.[60]

No campo jurídico, o debate incluiu a interpretação do art. 22, das possíveis interpretações das cortes europeias e de princípios gerais do direito da privacidade e proteção de dados bem como os limites e as falhas do GDPR para endereçar os problemas relacionados a transparência e *accountability* algorítmica em relação aos direitos de propriedade intelectual e ao segredo de negócio.

Nesse sentido, a questão sobre os algoritmos enquanto componentes de cadeias econômicas da economia da informação perpassa todo esse trabalho. Mas é importante destacar que tal problema envolve discussões em diversos campos dos quais não é possível extrair nenhuma síntese definitiva. São questões políticas e sociais em disputa e cujas soluções jurídicas encontram-se em construção.

Portanto, apesar de tratar de temas como a propriedade intelectual e os tipos de modelos algoritmos de inteligência artificial, as conclusões desta tese para esses problemas não devem ser consideradas como respostas definitivas, visto que algumas abordagens nessas áreas não foram analisadas neste estudo com a profundidade necessária, tais como as questões técnico-computacionais. No entanto, buscou-se, sempre que possível, apresentar as diferentes visões sobre algumas questões específicas.

Nesse sentido, este trabalho é construído a partir de uma escolha metodológica que engloba diferentes campos da disciplina jurídica de distintos ordenamentos e de abordagens multidisciplinares, com o foco na compreensão de uma questão específica sobre a existência de um direito. As principais fontes utilizadas na tese podem ser divididas da seguinte forma:

i) Em primeiro lugar, como condição essencial deste trabalho, partiu-se de uma análise das regulações pertinentes ao problema do direito à explicação, como o GDPR, a LGPD, os guias e documentos das autoridades nacionais de proteção de dados. Neste caso, diante do pouco tempo em que a LGPD entrou em vigor e da criação da ANPD, em muitos casos os materiais se restringem às regulações europeias. Neste sentido, mapearam-se as diferenças e

[60] A partir de 2016, as publicações sobre o tema explodiram. Uma pesquisa na base de dados da biblioteca digital da Association for Computing Machinery (ACM) pode fornecer uma dimensão desse processo. Do início da base de dados de periódicos disponíveis, até o ano de 2015, foram 413 resultados de trabalhos com a palavra "*explanation*" no título dos artigos. No curto período entre 2016 e 2021, esse número atingiu mais de 73 mil resultados.

similaridades dos dois ordenamentos quanto ao direito à explicação e as possíveis interpretações;

ii) Além das fontes legais, esta tese baseou-se na análise dos principais trabalhos que discutiram a existência ou não do direito à explicação no contexto Europeu e do Brasileiro desde a aprovação do GDPR e da LGPD, incluindo as visões favoráveis e contrárias a existência desse direito, como forma de elencar os principais argumentos e lacunas desses trabalhos;

iii) A resposta aos questionamentos sobre o tema dependeu do estudo sobre os princípios da proteção de dados, notadamente os relacionados à transparência, *accountability* e a autodeterminação informacional na doutrina de proteção de dados no contexto europeu e brasileiro, bem como dos trabalhos relacionados ao conceito de dados pessoais. Além desses trabalhos, entidades da sociedade civil que se debruçam sobre a discussão da transparência também tem realizado trabalhos sobre o tema que foram considerados no estudo;

iv) Outro elemento importante para a argumentação desta tese diz respeito à construção de uma síntese sobre o debate da transparência algorítmica na doutrina sobre a proteção de dados pessoais e nos trabalhos sobre ciência da computação. A preocupação com os processamentos de dados no contexto de *big data* são os elementos centrais por trás das iniciativas legislativas e regulatórias, sua compreensão contribui para definição do problema cujas legislações buscaram resolver. Essa síntese incluiu a tradição dos EUA. Esta abordagem do direito norte americano é relevante por dois motivos. O primeiro deles diz respeito à abordagem própria norte-americana sobre a temática de proteção de dados, outro deles diz respeito a própria contribuição que essa diferente abordagem pode fornecer a compreensão do problema e a interpretação dos dispositivos sobre as decisões automatizadas, possibilitando observar a questão da transparência nas decisões algorítmicas a partir da ótica do devido processo legal;

v) Para a compreensão do direito à explicação no direito brasileiro, fez-se necessário a análise de legislações anteriores à LGPD que disciplinaram direitos de transparência no contexto de tratamento de dados pessoais no contexto do direito do consumidor. A interpretação de alguns julgados de referência sobre os *scores* de crédito a partir do Código de Defesa do

Consumidor e da Lei do Cadastro Positivo fornecem bons elementos para a compreensão de como o acolhimento do Art. 20 da LGPD deve ocorrer na realidade brasileira;

vi) Além dessas abordagens, o trabalho se utiliza da análise de precedentes da justiça brasileira e europeia sobre decisões automatizadas, obrigações de transparência e direito à informação no contexto da proteção de dados. Para o caso Europeu, utilizou-se de bases de dados[61] que reúnem decisões de tribunais e autoridades nacionais europeias nos temas da proteção de dados e onde é possível segmentar as buscas com base nos artigos envolvidos nas decisões. Pesquisou-se, também, a base do Tribunal de Justiça da União Europeia, bem como o Tribunal Europeu de Direitos Humanos (TEDH) de Estrasburgo. Consideramos neste trabalho os casos em trâmite até março de 2021, de forma que é possível que existam novos casos. Já os precedentes analisados para tratar do caso brasileiro são julgados já amplamente debatidos pela literatura nacional sobre proteção de dados e direitos do consumidor.

No debate da transparência algorítmica, uma série de trabalhos propõe metodologias para implementação de explicações como mecanismo de transparência, incluindo trabalhos patrocinados ou mesmo realizados por entidades reguladoras. Importantes atores na academia e em organizações da sociedade civil propuseram modelos de relatórios de impacto, de auditorias, de testes e de outras medidas de mitigação de riscos em decisões automatizadas. Com base nesses trabalhos e na síntese interpretativa realizada pelo trabalho, foi possível construir uma proposta de instrumentalização do direito à explicação que compõe o último capítulo desta tese.

1.2 SÍNTESE DO CAPÍTULO

No Capítulo 1, explicitamos o problema, a hipótese, a lacuna e a originalidade do presente trabalho. Nos capítulos seguintes, partindo do problema apresentado, qual seja, a incerteza sobre a existência de um direito à explicação na Lei Geral de Proteção de Dados Pessoais do Brasil, buscaremos explorar a hipótese de que é

[61] GDPRHUB. Welcome to GDPRhub. Disponível em: https://gdprhub.eu/index.php. 2021. Acesso em: 21 jun. 2021.

possível reconhecer a existência de tal direito em nosso ordenamento, por vezes com escopo ainda mais amplo do que aquele previsto na legislação europeia. Como salientado, a relevância desta investigação reside na escassez de estudos no cenário brasileiro sobre o direito à explicação, bem como na originalidade do tema, uma vez que, conforme mapeado, inexistem trabalhos a nível de mestrado e doutorado no Brasil se debruçando especificamente sobre este assunto. Na seção 1.1.3 apresentamos ainda os conceitos centrais e balizadores de todo o trabalho, como os conceitos de dado pessoal, algoritmo e de decisão. Por fim, na seção 1.2, apresentamos o recorte e algumas considerações metodológicas do trabalho. Nos capítulos seguintes da tese buscaremos apresentar o debate acadêmico e hermenêutico em torno do direito à explicação, suas possíveis fontes e formas de reconhecimento, bem como suas limitações e desafios práticos de implementação. O Capítulo 2, apresentado a seguir, é o ponto de partida deste percurso argumentativo, possuindo o objetivo de discutir o que é o direito à explicação no cenário brasileiro e europeu, a partir de fontes positivadas, como a LGPD e a GDPR. Neste sentido, o capítulo buscará desenvolver o argumento de que é possível defender a existência desse direito no ordenamento brasileiro a partir de uma leitura sistemática da LGPD, das legislações setoriais de proteção de dados pessoais, da Constituição e da jurisprudência dos tribunais superiores. No âmbito da UE, apresentaremos de que forma ele pode ser extraído da leitura do texto da GDPR, bem como a leitura que vem sendo feita deste instrumento pela academia e pelas autoridades e órgãos de *enforcement* da UE. Além das fontes legais, o capítulo buscará apresentar ainda como o conceito de autodeterminação informativa, a cláusula geral dos direitos da personalidade e do devido processo informacional podem servir também como fundamento à existência do direito à explicação.

2 COMO É POSSÍVEL AFIRMAR A EXISTÊNCIA DE UM DIREITO À EXPLICAÇÃO?

Uma vez apresentados o desenho de pesquisa, o arcabouço metodológico e os pressupostos subjacentes a este trabalho, passaremos a reconstruir o debate transatlântico em torno da (in)existência de um direito à explicação, tomando como referência os trabalhos mais influentes sobre a questão, principalmente no contexto europeu, que compreendem o período que vai desde a debate legislativo até a aprovação da GDPR.

Na seção 2.1, especificamente, serão apresentadas, ainda de forma preliminar, as principais interpretações e posições em torno do reconhecimento de um direito à explicação a partir do direito positivado, sobretudo a partir das previsões contidas na GDPR. Essa discussão será posteriormente retomada no tópico 3.1.2.1, onde serão aprofundadas as diferentes interpretações dadas às disposições da GDPR por acadêmicos e pelas diretrizes do Article 29 Working Party.

Como veremos a seguir, não há um entendimento unânime acerca da efetiva existência de um direito à explicação na normativa europeia. Neste sentido, ainda que seja possível sustentar a existência de um direito à explicação a partir da interpretação sistemática e teleológica de diplomas como a LGPD e a GDPR, a seção 2.2 se propõe a dar um passo além ao discutir os possíveis fundamentos metajurídicos do direito à explicação, abordando formas de compreensão de um direito que vão além daquelas expressamente previstas na legislação.

Nesta seção, serão apresentados, preliminarmente, os principais elementos para a compreensão da problemática, quais sejam: o papel dos dados pessoais nas tecnologias de automação, como algoritmos e *machine learning*, por exemplo, os arts. 13, 14, 15 e 22 da GDPR, o *Recital* 71, e o art. 20 da LGPD, bem como as principais limitações e desafios associados ao direito à explicação. As limitações e desafios, por sua vez, serão objeto de discussão mais detalhada em capítulo próprio (Capítulo 4), onde serão aprofundados.

Um dos elementos centrais para a compreensão da problemática que aqui se apresenta consiste no papel desempenhado pelos dados pessoais nos sistemas automatizados. O debate sobre sistemas algorítmicos esbarra necessariamente no campo dos dados pessoais. Isso porque o substrato de operação de qualquer algoritmo são os dados, e é intrínseco à personalização de qualquer serviço voltado

para pessoas naturais, em qualquer nível, o uso de dados pessoais[62]. Regular o uso e o tratamento de dados pessoais é o principal objeto de leis de proteção de dados. Essas leis carregam efeitos diretos nas limitações e obrigações atinentes ao emprego de algoritmos. Como veremos adiante, são camadas distintas do mesmo processo, mas indissociáveis.

As leis de proteção de dados visam não somente proteger a privacidade, mas também garantir outros direitos fundamentais e liberdades individuais[63], que somente podem ser exercidos na sua completude caso seja garantido o uso adequado dos dados pessoais, entendidos como uma representação do indivíduo a partir de múltiplas facetas arbitrárias[64]. Assim, pode-se entender as leis de proteção de dados como um plexo regulatório que acaba por proteger outros direitos, uma vez que o exercício de liberdades, principalmente na esfera digital, passa pelo tratamento de dados pessoais[65]. Os contextos nos quais decisões automatizadas têm impactado no exercício e acesso a uma série de direitos fundamentais são variados e complexos[66], e necessariamente passam pelo uso adequado ou inadequado de dados pessoais. Todavia, a opacidade com a qual os dados pessoais são tratados impede que seus titulares tenham total compreensão de como seus direitos podem ser limitados e suas vidas impactadas[67] .

Tais contextos dão origem à necessidade de um novo construto normativo capaz de garantir proteção efetiva a seus titulares[68], e não somente proteger os dados pessoais em si. É necessário focar mais no impacto que o uso de dados pessoais

---

[62] FRAZÃO, A. Fundamentos da proteção dos dados pessoais: noções introdutórias para a compreensão da importância da Lei Geral de Proteção de Dados. *In*: TEPEDINO, Gustavo; FRAZÃO, Ana; OLIVA, Milena Donato. *Lei Geral de Proteção de Dados Pessoais e suas repercussões no Direito Brasileiro*. 2. ed. São Paulo: Thomson Reuters, 2020. p. 23-52.

[63] WACHTER, S.; MITTELSTADT, B.. A Right to Reasonable Inferences: Re-thinking Data Protection Law in the Age of Big Data and AI. *Columbia Business Law Review*, 2019, v. 2, 5 out. 2018. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3248829. Acesso em: 16 abr. 2021.

[64] CHENEY-LIPPOLD, J. *We are data*: algorithms and the making of our digital selves. New York: New York University Press, 2017.

[65] "A proteção de dados pode ser vista como a soma de um conjunto de direitos que configuram a cidadania do novo milênio." (RODOTÀ, S. *A vida na sociedade da vigilância*: a privacidade hoje. Tradução: Danilo Doneda e Luciana Cabral Doneda. Rio de Janeiro: Renovar, 2008. p. 17).

[66] O'NEIL, C. *Weapons of Math Destruction:* How Big Data Increases Inequality and Threatens Democracy. New York, NY: Crown Books, 2016. E-book.

[67] PASQUALE, F. *The Black box society:* the secret algorithms that control money and information. First Harvard University Press paperback edition ed. Cambridge, Massachusetts; London, England: Harvard University Press, 2015.

[68] ACCESS NOW; AMNESTY INTERNATIONAL. *The Toronto Declaration:* protecting the right to equality and non-discrimination in machine learning systems. Toronto, 2018. Disponível em: https://www.accessnow.org/cms/assets/uploads/2018/08/The-Toronto-Declaration_ENG_08-2018.pdf. Acesso em: 15 dez. 2020.

pode ter nas pessoas do que simplesmente na conformidade de tais usos a uma legislação vigente.[69] É a partir desta demanda que se entende o chamado *direito à explicação*, sobre o qual passaremos a tratar a seguir.

[69] Mayer-Schönberger, ao discorrer sobre a evolução das leis de proteção de dados pessoais na Europa, reconhece a crescente centralidade assumida pelo indivíduo nas recentes legislações sobre a matéria. De acordo com o autor, esse movimento pode ser observado sobretudo a partir da 2ª geração de leis de proteção de dados. A primeira geração surge ainda nas décadas de 1960 e 1970, num contexto caracterizado por um maior uso de dados por parte do Estado e grandes corporações, bem como pelo advento de novas tecnologias e práticas que possibilitaram sua agregação e tratamento em larga escala. Nessa fase, o foco das regulações estava mais voltado para o processamento e o fluxo de dados, não necessariamente para a tutela do indivíduo. A segunda geração, que compreende as leis originadas no final da década de 1970, pode ser caracterizada por uma reorientação do regime de proteção de dados: parte-se de um regime de regulação da tecnologia para um regime que dá mais foco à garantia das liberdades e direitos dos indivíduos. Neste período, vigora como paradigma a ideia de privacidade enquanto liberdade negativa. A partir da terceira geração, na década de 1980, há uma transição do paradigma da liberdade negativa para um paradigma baseado no reconhecimento da proteção de dados pessoais enquanto um direito autônomo e como liberdade positiva. Nesse momento, confere-se maior foco e força à ideia de *autodeterminação informacional*, ou seja, reconhece-se a capacidade do indivíduo decidir e participar ativamente das decisões envolvendo o fluxo de seus dados, movimento que sofre forte influxo do julgamento paradigmático do Tribunal Constitucional Federal da Alemanha no caso do recenseamento da população, em 1983. Ato contínuo, a partir da década de 1990, há a emergência de uma quarta geração de leis de proteção de dados, que conjugam a tutela individual a mecanismos de tutela coletiva. Tem-se, ainda, a coexistência de leis gerais com leis setoriais que as complementam, bem como o fortalecimento do aparato institucional de *enforcement*, com a disseminação de autoridades independentes. De acordo com Mayer-Schönberger, esse processo evolutivo não chegou ao fim, sendo possível que novas gerações sejam identificadas a partir das respostas legislativas dadas aos novos desafios e desenvolvimentos tecnológicos. Cf. MAYER-SCHOENBERGER, V. Generational development of data protection in Europe. *In*: AGRE, P. E.; ROTENBERG, M. (Eds.). *Technology and Privacy:* The New Landscape. Cambridge: MIT Press, 1998. p. 219-241. Apesar de possibilitarem uma tutela de escopo mais amplo, as recentes leis gerais de proteção de dados ainda apresentam algumas limitações que acabam deixando de fora de seu escopo relevantes situações e contextos de tratamento de dados. Tendo em vista os recentes desenvolvimentos tecnológicos e as novas práticas de processamento possibilitadas pelo Big Data e IA, alguns autores têm advogado por uma remodelação do atual regime de proteção de dados e pelo reconhecimento de novos direitos, que possuem em comum o objetivo de garantir uma maior proteção do indivíduo. Nesse sentido: "Privacy legislation has not kept up with these developments and remains based on concepts developed in the mainframe era. Thus, we need a new generation of privacy governance to cope with the implications of the new generation of online technologies, and to protect the new generation of technology users." [...] "All this calls for a fundamental reshuffle of the Current Framework; for the application of a legal regime attuned to a risk of harm continuum, rather than a dichotomy between personal and non-personal data, or private and public spheres; for a new approach to notice and choice emphasizing user awareness and understanding, rather than corporate disclaimers of liability; for an allocation of responsibility according to data use and risks to data subjects, rather than a formal dichotomy between controllers and processors; for a sensible balance between data retention needs and individuals' 'droit a` l'oubli'; and for cross border data transfers governed by accountability and ongoing responsibility, rather than arbitrary barriers and bureaucratic form filling." (TENE, Omer. Privacy: The new generations. *International Data Privacy Law*, v. 1, n. 1, p. 15-27, fev. 2011. Disponível em: https://academic.oup.com/idpl/article/1/1/15/759641. Acesso em; 21 jun. 2021. p. 15; 26-27; e WACHTER, S.; MITTELSTADT, B.. A Right to Reasonable Inferences: Re-thinking Data Protection Law in the Age of Big Data and AI. *Columbia Business Law Review*, 2019, v. 2, 5 out. 2018. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3248829. Acesso em: 16 abr. 2021).

2.1 O QUE É O DIREITO À EXPLICAÇÃO?

O direito à explicação pode ser entendido como o direito a receber informações suficientes e inteligíveis que permitam ao titular dos dados, e à sociedade, *entenderem, compreenderem,* a lógica, a forma, os critérios utilizados para tratar dados pessoais e, a partir dessa explicação, prever, antever os seus impactos[70], com o fim de evitar práticas discriminatórias, ilegítimas e indesejadas, que podem ter impacto no plano individual e coletivo, para que possam, ainda, ser desafiadas e revisadas por meio do exercício de direitos e do devido processo.

O direito à explicação seria regulado, no contexto nacional, em certa extensão, pela Lei Geral de Proteção de Dados do Brasil (Lei nº 13.709/2018), em conjunto com disposições contidas em algumas leis setoriais, como a Lei do Cadastro Positivo (Lei nº 12.414/2011) e o Código de Defesa do Consumidor (Lei nº 8.078/1990), e, no contexto da União Europeia, pelo Regulamento Geral de Proteção de Dados (GDPR, ou RGPD, na versão portuguesa). Esses diplomas, todavia, trazem consigo algumas limitações que podem ter impacto no efetivo exercício deste direito, como veremos ao longo deste trabalho.[71]

Como afirmado anteriormente, os principais elementos normativos para a compreensão do direito à explicação no regulamento europeu encontram-se nos arts. 13, 14, 15 e 22 da GDPR, bem como no seu Considerando 71. O art. 22(1) da GDPR consagra uma regra geral de proibição[72] a decisões tomadas exclusivamente com base no processamento automatizado de dados, incluindo a prática de *profiling*, que venha a produzir efeitos na esfera jurídica do titular dos dados ou que o afete significativamente de forma similar. De acordo com o texto, o controlador deverá se abster de levar a cabo a atividade de processamento descrita no art. 22(1), a menos que uma das exceções previstas no art. 22(2) sejam aplicáveis, quais sejam: (a) se a decisão for necessária para a celebração ou a execução de um contrato entre o titular

---

[70] WACHTER, S.; MITTELSTADT, B. A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI. *LawArXiv*, 12 out. 2018. Disponível em: https://doi.org/10.31228/osf.io/mu2kf. Acesso em: 27 maio 2020.

[71] MONTEIRO, R. L. *Existe um direito à explicação na Lei Geral de Proteção de Dados do Brasil?* Rio de Janeiro: Instituto Igarapé, dez. 2018. Artigo Estratégico nº 39. Disponível em: https://igarape.org.br/wp-content/uploads/2018/12/Existe-um-direito-a-explicacao-na-Lei-Geral-de-Protecao-de-Dados-no-Brasil.pdf. Acesso em: 21 jun. 2021.

[72] EUROPEAN DATA PROTECTION BOARD. *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 679/2016*. Bruxelas: European Commission, 2016. Disponível em: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053;

dos dados e o responsável pelo tratamento; (b) se a decisão restar autorizada pelo direito regional ou nacional do Estado-Membro ao qual o responsável estiver sujeito, observadas as garantias e liberdades do titular; e (c) se a decisão estiver autorizada pelo consentimento explícito do titular dos dados.

O art. 22(3), por sua vez, nas hipóteses das alíneas "a" e "c", confere ao titular sujeito à decisão automatizada o direito de ter salvaguardados seus direitos, liberdades e legítimos interesses, designadamente o direito de, pelo menos, obter intervenção humana por parte do responsável, manifestar o seu ponto de vista e contestar a decisão.

Como se pode observar, a GDPR não prevê, ao menos explicitamente e de modo vinculante, um "direito à explicação", por mais que imponha práticas e direitos de transparência e um direito de revisão. A existência de um direito à explicação tem sido sustentada por alguns autores com base em uma interpretação holística do texto do regulamento, a saber, de uma leitura sistemática dos arts. 13, 14, 15 e 22, bem como do Considerando 71, que apesar de não ser juridicamente vinculante, expressamente amplia as salvaguardas do art. 22(3) ao prever um "direito à explicação" no contexto de decisões automatizadas[73]. Os arts 13 e 14 encerram um conjunto de obrigações de transparência. O art. 15, por sua vez, consagra um direito de acesso. Em conjunto, esses três dispositivos preveem a necessidade de que sejam fornecidas informações significativas (*meaningful information*) sobre a existência de decisão automatizada, incluindo *profiling*.[74] O direito à explicação, assim, na GDPR, seria derivado dos direitos e garantias de não sujeição a decisões automatizadas (art. 22(1) e (3)), bem como dos deveres de notificação e informação dos controladores e do direito de acesso (arts. 13-15).

Conforme apontamos, não há, contudo, consenso no debate europeu acerca da efetiva existência de um direito à explicação vinculante, ou mesmo acerca de seu escopo e extensão.[75] Nesse sentido, há ao menos duas correntes no debate europeu

---

[73] Considerando 71: "In any case, such processing should be subject to suitable safeguards, which should include specific information to the data subject and the right to obtain human intervention, to express his or her point of view, *to obtain an explanation of the decision reached after such assessment and to challenge the decision*. Such measure should not concern a child." (UNIÃO EUROPEIA. *Regulamento Geral sobre a Proteção de Dados — Considerando 71*. Disponível em: https://gdpr-text.com/pt/read/recital-71/. Acesso em: 21 jun. 2021). (grifo nosso).

[74] SELBST, A. D.; POWLES, J. Meaningful Information and the Right to Explanation. *International Data Privacy Law*, v. 7, n. 4, p. 233-242, 2017. Disponível em: https://papers.ssrn.com/abstract=3039125. Acesso em: 27 maio 2020.

[75] CASEY, B.; FARHANGI, A.; VOGL, R. Rethinking Explainable Machines: The GDPR's "Right to Explanation" Debate and the Rise of Algorithmic Audits in Enterprise. In: *Berkeley Technology Law Journal*, v. 34, p. 145-189, 2018. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3143325. Acesso em: 19 maio 2019. p. 145-189.

acerca do direito à explicação.[76]

A primeira é representada por Wachter, Mittelstadt e Floridi[77], que sustentam não haver um efetivo "direito à explicação" na GDPR. De acordo com os autores, há dois tipos de explicações possíveis para decisões automatizadas baseadas em Inteligência Artificial[78]: (i) explicações quanto ao funcionamento do algoritmo e (ii) explicações quanto aos fundamentos da decisão conferida. Em razão de um conjunto de fatores técnicos e legais (e.g. direitos de propriedade intelectual, dificuldades quanto ao fornecimento de explicações acessíveis (*meaningful information*) quanto aos aspectos técnicos do sistema, a jurisprudência legada da Diretiva 95/46 e uma interpretação restritiva dos arts. 13-15 da GDPR), os autores sustentam haver apenas um direito limitado à explicação quanto ao funcionamento do algoritmo, o que eles chamaram de "direito à informação".

De outro lado, em resposta à Wachter *et al.*[79], Julia Powles e Andrew Selbst[80] afirmam que, embora a GDPR não preveja expressamente um "direito à explicação", ele não deve ser entendido como sendo *inexistente* ou *ilusório*. De acordo com Selbst e Powles, o direito à explicação seria derivado do direito a uma *informação significativa* (*meaningful information*), contido nos arts. 13, 14 e 15 da GDPR.

Em suma, os que defendem a existência afirmam categoricamente que a GDPR, ao estabelecer direitos de informação sobre a lógica de processos de decisões automatizadas[81], confere claramente o direito à explicação, e esse deve ser interpretado de modo a permitir ao titular dos dados o exercício de seus direitos previstos na própria regulamentação e no ordenamento jurídico[82]. Por outro lado, os que se opõem argumentam que a ausência do termo "explicação" no texto vinculativo da GDPR não permitiria afirmar categoricamente a existência de tal direito na

---

[76] INSTITUTO DE REFERÊNCIA EM INTERNET E SOCIEDADE. Right to Explanation and Artificial Intelligence. *IRIS BH*, 17 jun. 2019. Disponível em: https://irisbh.com.br/en/right-to-explanation-and-artificial intelligence-brief-considerations-on-the-european-debate/. Acesso em: 7 nov. 2020.

[77] WACHTER, S.; MITTELSTADT, B.; FLORIDI, L. Why a Right to Explanation of Automated Decision Making Does Exist in the General Data Protection Regulation. *In: International Data Privacy Law*, vol. 7, n. 2, maio 2017, p. 76–99. Disponível em: https://academic.oup.com/idpl/article/7/2/76/3860948. Acesso em: 6 nov. 2020.

[78] A presente pesquisa possui como foco decisões automatizadas baseadas em Inteligência Artificial e *Machine Learning*.

[79] WACHTER; MITTELSTADT; FLORIDI, op. cit.

[80] SELBST, A. D.; POWLES, J. Meaningful Information and the Right to Explanation. *International Data Privacy Law*, v. 7, n. 4, p. 233-242, 2017. Disponível em: https://papers.ssrn.com/abstract=3039125. Acesso em: 27 maio 2020.

[81] Os arts. 13 e 14 da GDPR garantem o direito a "informações úteis relativas à lógica subjacente."

[82] SELBST; POWLES, op. cit.

amplitude semântica do termo[83].

Para além das posições que argumentam pela existência ou inexistência de um direito à explicação na GDPR, há autores que apontam para o fato de a discussão ter caminhado para um foco excessivamente reducionista e ter desconsiderado aspectos e salvaguardas mais amplos trazidos pela texto do Regulamento, como o regime de *accountability*, o paradigma *ex ante* de regulação baseado no risco e na elaboração prévia de *Data Protection Impact Assessments* (DPIAs) quando estiver em jogo atividades de processamento de alto risco e o conjunto de princípios de transparência e *Data Protection by Design* (DPbD).[84]

Para além de um mero direito individual, baseado numa lógica de requisições *ex post* demandando o exercício do direito à explicação, o paradigma *ex ante* seria capaz de garantir salvaguardas mais amplas, que alcançariam o próprio *design* e a implementação de decisões automatizadas. Caso não sejam suficientes para justificar a desnecessidade ou impertinência do reconhecimento de um direito à explicação, a consideração desses outros elementos normativos no debate serviria ao menos para tornar esse direito significativamente mais robusto.

Neste sentido, ao discorrer sobre a transição de um paradigma individual, baseado em requisições pontuais e *ex post* acerca do direito à explicação, para um paradigma baseado em Data Protection by Design (DPbD), Casey *et al.*[85] argumentam que os benefícios são inúmeros. Em primeiro lugar, porque a transparência resultante do atendimento de um direito à explicação pode não ser necessariamente útil para o indivíduo, que pode ter dificuldades para instrumentalizá-la e tirar proveito da informação que lhe fora conferida[86]. Segundo que um regime focado em DPbD garantiria uma maior flexibilidade às empresas para implementar sistemas de aprendizado de máquina, sem restringir demasiadamente a inovação[87]. Terceiro, a implementação de auditorias e de princípios de DPbD tende a ser mais vantajosa quando comparada a um paradigma calcado no exercício individual do direito à

[83] WACHTER, S.; MITTELSTADT, B.; RUSSELL, C. Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR. *SSRN Electronic Journal*, v. 31, n. 2, 2017. Disponível em: https://www.ssrn.com/abstract=3063289. Acesso em: 27 maio. 2020.
[84] CASEY, B.; FARHANGI, A.; VOGL, R. Rethinking Explainable Machines: The GDPR's "Right to Explanation" Debate and the Rise of Algorithmic Audits in Enterprise. In: *Berkeley Technology Law Journal*, v. 34, p. 145-189, 2018. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3143325. Acesso em: 19 maio 2019. p. 180.
[85] Ibidem, p. 181-184.
[86] Ibidem, p. 181.
[87] Ibidem, p. 182.

explicação, pois são capazes de promover uma transparência mais ampla sobre os sistemas algorítmicos, possibilitando que a mídia, ONGs, especialistas e outros atores auditem os sistemas, e não apenas os indivíduos. Ademais, essas metodologias também têm sido documentadamente bem-sucedidas na identificação de vieses discriminatórios nos sistemas algorítmicos.[88]

Ainda em 2017, Edwards e Veale escreveram o artigo "Slave to the Algorithm: Why a Right to an Explanation is Probably Not the Remedy You Are Looking For"[89], cujo argumento central vai além da existência ou não de um direito à explicação e foca na ideia de que um direito desta natureza dificilmente representaria um remédio adequado para responder a todos os riscos apresentados por algoritmos e tecnologias correlatas. O artigo afirma, em primeiro lugar, que o texto da GDPR é restritivo, incerto e até paradoxal sobre quando um direito de explicação pode ser acionado. Depois, sobre a questão de aplicabilidade prática, cogita a inviabilidade de que o tipo de explicação almejada pelos seus defensores seja contemplada, ou até viável, pelos tipos de explicação de *machine learning* que cientistas da computação têm desenvolvido.

Os autores receiam que a busca pelo direito à explicação crie uma "falácia da transparência"[90] e argumentam que outros direitos, como o direito ao esquecimento e portabilidade de dados e outros elementos, como *Privacy by Design*, *Data Protection Impact Assessments*, certificações e selos de privacidade podem ser um caminho melhor para tornar algoritmos mais responsáveis, explicáveis e centrados na experiência humana.

Como introdução sobre o papel crescente do algoritmo na sociedade, os autores afirmam que os indivíduos se tornaram "escravos dos algoritmos" e fazem referência ao termo de Frank Pasquale da "sociedade *black box*". Analisando os principais problemas gestados neste contexto, com destaque para o avanço das técnicas de *machine learning*, os autores destacam i) discriminação e vieses injustos; ii) riscos à privacidade, diante do fato de que estas técnicas baseiam-se na ideia de

---

[88] CASEY, B.; FARHANGI, A.; VOGL, R. Rethinking Explainable Machines: The GDPR's "Right to Explanation" Debate and the Rise of Algorithmic Audits in Enterprise. In: *Berkeley Technology Law Journal*, v. 34, p. 145-189, 2018. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3143325. Acesso em: 19 maio 2019. p. 181-184.
[89] EDWARDS, L.; VEALE, M. Slave to the Algorithm: Why a Right to an Explanation Is Probably Not the Remedy You Are Looking for. *Duke Law & Technology Review*, v. 16, p. 18-84, 2017. p. 18.
[90] HOOD, C.; HEAD, D. *Transparency*: The Key to Better Governance? Oxford: Oxford University Press, British Academy, 2006.

dar novos propósitos para dados, o que contraria o princípio da finalidade na disciplina da proteção de dados pessoais; iii) opacidade, ponto em que os autores discorrem sobre a singularidade da transparência enquanto valor histórico insculpido nas legislações de proteção de dados pessoais e como vem sendo elevada ao centro do debate sobre *accountability* para algoritmos e *machine learning.*

Sobre o direito à explicação, Edwards e Veale alegam que a noção de um direito à explicação já existia na Diretiva desde 1995, e foi incorporada à GDPR, mas com uma série de limitações, tais como exceções à Propriedade Intelectual e segredo de negócio, limitação do escopo a decisões tomadas "exclusivamente" por sistemas automatizados e que produzam efeitos legais ou "significativos", o *timing* da explicação, a presença dos elementos mais "fortes" do direito à explicação nos Considerandos e *guidelines* do Working Party 29, e não no texto em si do Regulamento e, por fim, as dificuldades práticas. Dessa forma, pode-se dizer que, quanto ao debate primordial sobre a existência do direito à explicação, os autores filiam-se à corrente de Wachter *et al.*

A partir disso, o *paper* não só questiona a existência do direito, mas também se este seria de fato o melhor direito a se buscar para garantir a *accountability* dos algoritmos e decisões automatizadas. Um primeiro problema identificado pelos autores é que os remédios tradicionais de proteção de dados pessoais, aos quais associa este modelo de direito à explicação, são voltados para o indivíduo, já que o sistema deriva de um paradigma de direitos humanos, mas os riscos envolvidos na discussão normalmente são associados a grupos.

Também apontam que, em muitos casos, o titular de dados pessoais estará menos interessado em uma explicação sobre uma decisão e mais em que a decisão em questão sequer ocorra ou que, diante dela, haja recursos para que seja revertida[91]. Por fim, ainda que não se confirme esta assertiva e os indivíduos de fato tenham interesse em obter informações, os autores vislumbram um obstáculo significativo na falta de tempo, recursos e conhecimento da maioria dos indivíduos para que consigam de fato absorver qualquer explicação que o sistema lhes ofereça. Assim, os autores comparam o recurso a um direito individual à explicação ao *status* atual que atribuem

[91] Ao desenvolver este argumento, os autores fazem referência ao *caso Costeja v. Google*, emblemático para a discussão sobre direito ao esquecimento, e apontam que a demanda do titular de dados pessoais, nesse caso, era voltada para o exercício de um direito de ação, com o objetivo de suprimir a circulação destes dados, e em nenhum momento obter uma explicação sobre eles.

ao consentimento num mundo virtual — o de uma “falácia da transparência”.

Dentre as outras dificuldades apresentadas, o artigo opta por focar na viabilidade técnica das explicações, isto é, em uma fase de eventual implementação. Porque o *machine learning* opera por meio da construção de modelos alimentados com dados de *input* para propósitos preditivos e porque a sua concepção é indutiva e tem como prioridade a performance e não a interpretabilidade, conforme estes modelos se tornam mais complexos, eles podem ser incompreensíveis para humanos. O significado do “aprendizado” no *machine learning* diz respeito à capacidade de aprimoramento do modelo, o que é atingido e mensurado não pela investigação da sua estrutura interna, mas pela análise do seu comportamento por meio de métricas de performance.

Nesse contexto, ainda, os autores observam o *trade-off* entre performance e explicabilidade[92], na medida em que sistemas com mais variáveis (mais difíceis de explicar do ponto de vista interno) tendem a operar melhor. Somados estes percalços de ordem técnica com as observações acerca do interesse do titular de dados, os autores levantam um possível questionamento acerca do custo de explicações focadas no indivíduo *versus* seus benefícios.

Diante de todas estas ressalvas ao direito de explicação como solução para a questão da transparência e *accountability* de sistemas de *machine learning*, o *paper* levanta algumas possibilidades de caminhos para o debate — a primeira volta um olhar para outros direitos garantidos pelo Regulamento, como o direito ao esquecimento[93] e a portabilidade de dados[94].

---

[92] O livro “For a meaningful artificial intelligence: towards a french and european strategy”, organizado por Cedric Villani, dá grande destaque também à questão do *trade-off* no capítulo 5, denominado “What are the ethics of AI?”. (VILLANI, C. *For a Meaningful Artificial Intelligence:* Towards a French and European Strategy. [*S. l.*]: European Commission, 2018. Disponível em: https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf).

[93] “Art. 17 GDPR — Right to erasure (‘right to be forgotten’)
1. The data subject shall have the right to obtain from the controller the erasure of personal data concerning him or her without undue delay and the controller shall have the obligation to erase personal data without undue delay where one of the following grounds applies: [...].” (UNIÃO EUROPEIA. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance)*. Official Journal of the European Union, 4 maio 2016. Disponível em; https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng. Acesso em: 27 ago. 2021).

[94] “Art. 20. Right to data portability
1. The data subject shall have the right to receive the personal data concerning him or her, which he or she has provided to a controller, in a structured, commonly used and machine-readable format and have the right to transmit those data to another controller without hindrance from the controller to which the

Quanto ao direito ao esquecimento, que pode ser invocado se os dados pessoais forem usados para fins de *profiling* e tiverem sido coletados legalmente, mas sem consentimento, além de depender de algumas bases legais específicas e possuir várias exceções, também está em aberto se inferências, elemento central do debate aqui travado, podem ser requeridas/deletadas. Sobre a portabilidade, por sua vez, os autores introduzem a possibilidade de que os usuários protejam seus dados pessoais tomando posse deles e utilizando *Personal Data Containers* (PDCs), espécies de "depósitos" para compartimentalização de dados pessoais. Os questionamentos levantados em relação a esta abordagem são dois: primeiro, em uma economia movida a base de dados, o isolamento do indivíduo pode não ser algo desejável; além disso, o artigo referente à portabilidade de dados só se aplica a dados que o titular e dados obtidos com consentimento, não outras bases legais.

Discutidos os prós e contras dessas novas possibilidades, os autores partem para a abordagem que elencam como a mais atrativa diante do conjunto de desafios e oportunidades para o avanço de práticas de transparência e garantia de *accountability* em um mundo no qual os indivíduos se tornaram "escravos dos algoritmos": a abordagem sistêmica.

O principal problema do foco individual da regulação tradicional de proteção de dados pessoais, segundo os autores, é justamente que ela gera um peso sobre o indivíduo, que pode não ser capaz ou não ter interesse de carregar. Uma das diversas formas de balancear essa dinâmica se dá por meio dos mecanismos de supervisão das autoridades nacionais, mas os autores reiteradamente sugerem no artigo que o seu papel de árbitro, por si só, não seria suficiente, além do fato de sofrerem com limitações orçamentárias.

Outras exigências, estruturadas no âmbito do Regulamento e que se relacionam com uma visão sistêmica de *accountability* seriam mais promissoras: *Privacy by Design* e a obrigação, em determinadas hipóteses, de indicação de um *Data Protection Officer* (DPO) e produção de *Data Protection Impact Assessments (DPIAs)* ou Relatórios de Impacto sobre a Proteção de Dados Pessoais. Esses

---

personal data have been provided […]. (UNIÃO EUROPEIA. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance)*. Official Journal of the European Union, 4 maio 2016. Disponível em: https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng. Acesso em: 27 ago. 2021).

últimos, cuja compulsoriedade está atrelada ao risco do tratamento de dados pessoais e engatilha análises pelas autoridades nacionais, possuiria um potencial significativo de limitar práticas preditivas, segundo os autores.

Medidas voluntárias e colaborativas, como certificações e selos de qualidade, que podem ser setorializados, também são recebidas pelos autores com otimismo e todo esse conjunto de medidas sistêmicas ofertadas pela regulação europeia seria, de acordo com eles, uma alternativa para concretizar as aspirações de um “devido processo legal substantivo” ou “devido processo legal algorítmico", defendido por parte da doutrina estadunidense, como Citron, Pasquale, Crawford e Schultz, que não dispõe de uma figura tal qual o direito à explicação para direcionar suas aspirações sobre transparência e *accountability.*

Em suma, Edwards e Veale propõem uma atenção maior da literatura para duas questões: (i) a construção de melhores sistemas de *machine learning* desde sua concepção, a fim de endereçar as dificuldades técnicas de interpretabilidade destes sistemas e (ii) o empoderamento de indivíduos e entidades da sociedade civil para que revisem e avaliem a precisão, equidade e integridade dos sistemas como um todo, a partir dos instrumentos oferecidos pelo Regulamento europeu e não apenas desafiem decisões individuais.

Em 2018, duas novas contribuições ao debate sobre o direito à explicação sugeriram abordagens inovadoras e possivelmente negligenciadas na discussão focada na existência deste direito na GDPR. O primeiro deles, “Rethinking Explainable Machines: The GDPR's 'Right to Explanation' Debate and the Rise of Algorithmic Audits in Enterprise”[95] sustenta que, no ímpeto de encontrar um novo “direito revolucionário” na GDPR, os estudiosos do tema negligenciaram o que eles consideram o mais revolucionário do Regulamento — os enormes poderes de *enforcement* conferidos às autoridades nacionais de proteção de dados nos capítulos 6 e 8 e, por consequência, seus poderes interpretativos *de facto* sobre o direito à explicação. O texto defende que o verdadeiro poder desse direito deriva da sua dinâmica quando combinado com auditorias de algoritmos e “proteção de dados *by design”*.

[95] CASEY, B.; FARHANGI, A.; VOGL, R. Rethinking Explainable Machines: The GDPR’s “Right to Explanation” Debate and the Rise of Algorithmic Audits in Enterprise. In: *Berkeley Technology Law Journal*, v. 34, p. 145-189, 2018. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3143325. Acesso em: 19 maio 2019.

Introdutoriamente, o artigo retoma o direito à explicação no âmbito da GDPR e seus contornos ainda incertos, o que estaria na gênese de toda a discussão acadêmica em torno de sua existência, ou não, e seu alcance. O artigo defende que, em meio a este debate, o papel de *enforcement* das autoridades nacionais ficou esquecido. Diferente do ambiente regulatório da Diretiva que substitui, a GDPR confere às autoridades nacionais europeias poderes investigatórios alargados, uma "caixa de ferramentas" corretiva e a possibilidade de imposição de sanções pecuniárias cujos limites máximos são milhares de vezes maiores do que os estabelecidos previamente. Esses poderes, segundo os autores, convertem-se em poderes interpretativos e, se analisada por este prisma a problemática, muitas das discordâncias em torno do direito à explicação podem se tornar irrelevantes. A pergunta central do artigo, então, é: "O que aqueles de fato encarregados em fazer cumprir o referido direito acham que ele implica?"

Partindo de uma retomada do debate mobilizado ao redor da existência do direito de explicação, que dividem em três fases — o argumento original (Goodman e Flaxman), a resposta (Wachter *et al.*) e a réplica (Sebst e Powels) — os autores identificam falhas significativas em todas as três. No caso de Goodman e Flaxman, apontam certa superficialidade na análise e abordagem excessivamente técnica.

Quanto ao argumento pela inexistência do direito à explicação, Casey *et al* criticam ter sido "altamente seletivo" e ignorado pontos relevantes dos arts. 13, 14, 15 e 22, especialmente o significado do termo "significativa" em "informação significativa sobre a lógica envolvida nas decisões automatizadas". O artigo de Wachter et al também teria se preocupado pouco com as competências administrativas alargadas e a nova força (inclusive interpretativa) da GDPR. Por fim, os autores teriam ainda se contradito e proposto a existência de um "direito de ser informado", sem deixar clara a diferença substantiva entre ele e o direito à explicação. No que se refere ao redirecionamento da discussão operado por Sebst e Powles, os autores consideram-no salutar, bem como uma resposta satisfatória, mas alertam que tal contribuição foi escrita antes da maior parte das interpretações por parte das autoridades de dados mais influentes do bloco, acerca destas mesmas controvérsias, serem lançadas, o que tornaria a contribuição relativamente desatualizada.

O traço em comum entre todas as três abordagens, entretanto, seria terem passado ao largo do debate que Casey *et al*. consideram o mais importante — sobre os poderes das autoridades de proteção de dados pessoais. Além da evidente

diferença quanto à força das sanções respaldadas pelo novo Regulamento, a mais expressiva mudança trazida pela GDPR diria respeito à própria natureza da norma. Enquanto Diretivas criam regras gerais que precisam ser incorporadas pelas legislações de cada país conforme considerem apropriado, um Regulamento é uma lei única e diretamente aplicável sobre todos os Estados-membros.[96]

Considerados estes pontos, o artigo salienta dois Capítulos do Regulamento, em especial. O Capítulo 6 prevê

> [...] a indicação, por cada Estado-membro, de uma ou mais autoridades públicas independentes com poderes investigatórios, consultivos e corretivos, que devem garantir a "aplicação consistente" da GDPR, e incluem , dentre outras coisas, a capacidade de "obter [...] acesso a todos os dados pessoais [pertencentes a uma empresa] e a todas as informações necessárias para o desempenho de tarefas [de investigação] '(2)' realizar investigações sob a forma de auditorias de proteção de dados '(3)' emitir alertas [ou] reprimendas a [empresas], '(4)' impor uma limitação temporária ou definitiva [contra empresas] incluindo uma proibição de processamento", e '(5)' ordenar a suspensão de fluxos de dados para um destinatário em um país terceiro ou a uma organização internacional.[97]

O Capítulo 8, de forma complementar, estabelece as sanções para descumprimento das previsões do Regulamento. Mais relevante do que os altos valores, para os autores, é o foco nos *efeitos* das violações e na sensibilidade dos dados envolvidos, bem como a consideração sobre os esforços proativos das entidades, o que tende a estimular uma governança cooperativa saudável.

Como consequência lógica destes poderes de *enforcement* estaria uma busca ativa das empresas e entidades pelas orientações das autoridades nacionais a fim de assegurar *compliance* com a GDPR, motivo pelo qual os autores buscaram nos exemplos da autoridade inglesa[98] ("ICO") sinalizações interpretativas sobre o direito à

---

[96] CORBETT-DAVIES, S. *et al.* Algorithmic decision making and the cost of fairness. *ArXiv*, 2017. Disponível em: http://arxiv.org/abs/1701.08230. Acesso em: 14 set. 2020.

[97] UNIÃO EUROPEIA. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance)*. Official Journal of the European Union, 4 maio 2016. Disponível em; https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng. Acesso em: 27 ago. 2021.

[98] Justificam tal escolha por uma questão prática de eleger uma autoridade ao invés de 28 e também por ser um exemplo emblemático de país que, a despeito do Brexit, quer manter o livre fluxo transnacional de dados com a Europa por meio do *compliance* com a GDPR. Cabe pontuar que o principal documento considerado pelos autores para a redação do artigo foi o "*Guide to the UK General Data Protection Regulation (UK GDPR*)", e que após a publicação do paper o ICO produziu várias outras orientações sobre o direito à explicação no contexto de decisões automatizadas. Neste sentido, os

explicação. Os achados são resumidos da seguinte maneira: a ICO define que empresas e entidades que tratem dados pessoais que se enquadrem nas hipóteses do art. 22 do Regulamento assegurem que

> i) os titulares sejam informados acerca do tratamento; ii) sejam introduzidas formas simples de solicitar intervenção humana ou contestar uma decisão; iii) sejam realizadas verificações regulares para garantir que seus sistemas estão funcionando conforme o pretendido.[99]

A autoridade também requer expressamente que, sempre que se aplicar o art. 22, seja realizado um *Data Protection Impact Assessment* (DPIA) e, mesmo em situações fora do escopo do dispositivo, recomenda energicamente a realização de DPIAs como parte de uma ferramenta de *compliance* mais ampla baseada nos mesmos princípios de *data protection by design.*

Estas observações levam à constatação, pelos autores, de que o direito à explicação não é lido apenas como um remédio invocado por particulares, mas como uma forma mais geral de *oversight* com implicações profundas sobre "[...] *design*, prototipagem, testes de campo e implantação de sistemas de processamento de dados."[100] Dessa forma, é possível afirmar que o direito à explicação, por mais que seja um direito individual, tem consequências supraindividuais para que seja possível a sua viabilização. Ainda que o direito à explicação não venha a requerer uma abertura completa da caixa preta, ponto sobre o qual os autores não se debruçam detidamente, as orientações de autoridades europeias indicam requerer uma avaliação dos interesses de *stakeholders* relevantes, um entendimento sobre o processamento geral de seus sistemas e o desenvolvimento de práticas de documentação e justificação de determinadas características destes sistemas. Nesse sentido, para Casey *et al*, mais

próprios autores destacam que a autoridade concebe o documento como um instrumento vivo, sujeito a alterações e revisões de forma contínua. Uma importante iniciativa da autoridade britânica, em conjunto com o The Alan Turing Institute, após a publicação do *paper* foi o lançamento de uma consulta pública e de *guidelines* sobre a explicabilidade de decisões produzidas por meio de IA. Cf. INFORMATION COMMISSIONER'S OFFICE; THE ALAN TURING INSTITUTE. ICO and Turing consultation on Explaining AI decisions guidance. *ICO*, 24 jan. 2020. Disponível em: https://ico.org.uk/about-the-ico/ico-and-stakeholder-consultations/ico-and-the-turing-consultation-on-explaining-ai-decisions-guidance/. Acesso em: 16 abr. 2021.

[99] UNIÃO EUROPEIA. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance)*. Official Journal of the European Union, 4 maio 2016. Disponível em; https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng. Acesso em: 27 ago. 2021.

[100] Casey et. al, op. cit.

do que um direito *ex post* à explicação, um sistema *ex ante* de avaliação de riscos deve se tornar a norma para sistemas algorítmicos, na leitura apresentada pelo artigo.

Adiante, as vantagens de mudar o foco do debate sobre o direito à explicação seriam as seguintes: o primeiro é que o foco na transparência interna de sistemas pode colocar um peso excessivo sobre os indivíduos, que devem buscar e interpretar as informações (o que se convencionou denominar "falácia da transparência"). A provisão de explicações básicas diante de demandas individuais pode, ainda, dissimular as razões de empresas cujos sistemas de tratamento de dados sejam enviesados de outras formas. Além disso, em muitos casos, especialmente de sistemas de *machine learning* altamente complexos, o custo para obter uma explicação pode ser muito alto e superar o benefício dessa explicação a nível individual; por fim, o que é considerado mais importante é que auditorias de sistema do tipo vislumbrado pelos DPIAs já têm um histórico positivo na detecção e combate de discriminação em algoritmos opacos e têm a vantagem de permitir a interação com entidades externas, muitas vezes com mais recursos e conhecimento do que indivíduos isolados.

Na parte final do texto, os autores argumentam que, a despeito da crise que vem sofrendo, a União Europeia ainda tem um poder e uma influência imensas em termos de exportação de modelos e parâmetros legais para o resto do mundo, em um processo de "globalização regulatória unilateral". Esse fenômeno costuma ser chamado de "fenômeno Bruxelas". Um exemplo de fenômeno Bruxelas é justamente a Diretiva e como seus parâmetros se espalharam para o mundo inteiro, o que se explica facilmente como medida de mitigação de riscos de dificuldades comerciais com o maior bloco comercial do planeta. A GDPR estende essa lógica por meio de exigência expressa de adequação para assegurar presença no livro fluxo transnacional de dados. Embora se trate de um fenômeno observado no nível de Estados, parte da lógica de que o efeito Bruxelas tem se estendido também para entidades privadas, se adequando à GDPR ao redor do mundo.

Não obstante as divergências existentes, há consenso no debate europeu[101] e entre ambas as correntes apresentadas de que a redação da GDPR, quanto ao regime aplicável às decisões automatizadas, é dúbia e deixa dúvidas quanto à

[101] PRIVACY INTERNATIONAL. *Data is Power: Profiling and Automated Decision-Making in GDPR*. PI, 9 abr. 2017. Disponível em: https://privacyinternational.org/sites/default/files/2018-04/Data%20Is%20Power-Profiling%20and%20Automated%20Decision-Making%20in%20GDPR.pdf. Acesso em: 7 nov. 2020. p. 10.

existência e quanto ao escopo de um possível direito à explicação. Levantam-se ainda questões quanto à exequibilidade de um direito à explicação amplo, tendo em vista a dificuldade de se traduzir questões técnicas de funcionamento do sistema para o titular dos dados afetados pela decisão, de modo a garantir-lhe condições para a contestação da mesma[102].

É nesse contexto que emerge uma das principais limitações e desafios ao direito à explicação: como garantir que explicações sejam compreensíveis e úteis aos seus destinatários, de modo a assegurar que estes possam delas extrair proveito para contestar as decisões que afetem significativamente sua esfera de direitos?

Quanto a este aspecto, muito tem se discutido acerca de qual seria o formato ideal e quais os atributos necessários para que uma explicação possa ser considerada efetiva ou satisfatória. Conforme veremos no Capítulo 4, o debate acerca do modelo ideal de explicação deve considerar, necessariamente, os inúmeros desafios e limitações existentes, como os desafios relacionados à própria complexidade de sistemas algorítmicos, os limites da capacidade de compreensão de aspectos técnicos por parte dos titulares, abstratamente considerado como um usuário *médio*, sem *expertise* técnica, e limitações jurídicas associadas aos segredos de negócio e direitos de propriedade intelectual, por exemplo, que colocam significativas restrições à forma e à extensão de um direito à explicação.

Diferentemente da GDPR, o art. 20 da LGPD[103] não contém uma proibição geral às decisões automatizadas, apenas garantindo ao titular o direito de se opor caso elas venham a ser operacionalizadas e de obter do controlador informações *claras* e *adequadas* a respeito dos critérios e dos procedimentos utilizados. Não obstante, neste trabalho, conforme veremos na seção 3.2.2, a hipótese a ser desenvolvida é a de que tal direito, no contexto da LGPD, não apenas existe, como a forma como foi desenvolvido "[...] lhe dá um escopo de aplicação muito mais amplo do que aquele do contexto europeu."[104] Todavia, como acima apontado, a sua

---

[102] EDWARDS, L.; VEALE, M. Slave to the Algorithm: Why a Right to an Explanation Is Probably Not the Remedy You Are Looking for. *Duke Law & Technology Review*, v. 16, p. 18-84, 2017. Disponível em: https://osf.io/preprints/lawarxiv/97upg/. Acesso em: 7 nov. 2020.

[103] BRASIL. *Lei nº 13.709, de 14 de agosto de 2018*. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF: Presidência da República, 14 ago. 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015- 2018/2018/lei/L13709.htm. Acesso em: 07. nov. 2020.

[104] MONTEIRO, R. L. *Existe um direito à explicação na Lei Geral de Proteção de Dados do Brasil?* Rio de Janeiro: Instituto Igarapé, dez. 2018. Artigo Estratégico nº 39. Disponível em: https://igarape.org.br/wp-content/uploads/2018/12/Existe-um-direito-a-explicacao-na-Lei-Geral-de-Protecao-de-Dados-no-Brasil.pdf. Acesso em: 21 jun. 2021. p. 5. "Assim como na Lei europeia, o direito à explicação previsto no caso brasileiro pode encontrar algumas limitações, como a manutenção dos

implementação prática encontra diversos obstáculos, como os limites estabelecidos por regras de segredo de negócio e propriedade intelectual, em paralelo à complexidade cada vez maior dos algoritmos e sua opacidade quase que natural.[105] Essa dificuldade de implementação pode dificultar a garantia de exercício de outros direitos individuais e coletivos, exacerbar a assimetria de informação e, por consequência, a assimetria de poder, além de complexificar a demonstração de eventual responsabilidade por práticas inadequadas e abusivas no tratamento de dados pessoais.

Portanto, neste trabalho, partimos de uma preocupação em abordar e analisar tais potenciais insuficiências e dificuldades em garantir o direito à explicação, culminando em propostas capazes de saná-las. Mas, para isso, é necessário analisar alguns pressupostos.

2.1.1 O que são decisões automatizadas?

Já discutimos na seção 1.1.3 o conceito de decisão e como tal definição aceita diversas interpretações. As decisões foram definidas como um resultado dentre outros possíveis, a partir de um processo de avaliação e ponderação de um determinado conjunto de critérios. Mencionamos a discussão sobre a possibilidade de máquinas realizarem decisões. A discussão está em aberto, mas feitas as devidas ressalvas, para o objeto do presente trabalho, que é a compreensão de um possível direito à explicação de decisões tomadas com base no processamento automatizado de dados, faz-se necessário construir um conceito de decisão automatizada. Considerando que já apresentamos um conceito de decisão, resta-nos apresentar um conceito de automação, o qual buscaremos desenvolver ao longo dos próximos parágrafos.

A *International Society of Automation* define automação como "[...] a criação e aplicação de tecnologias para monitorar e controlar a produção e distribuição de bens

segredos industriais dos responsáveis pelo tratamento. Porém, o regulamento europeu impõe mais restrições do que a Lei brasileira, principalmente por não incluir o caso dos dados anonimizados e por limitar o direito de oposição quando a base legal para tratamento dos dados for o consentimento explícito ou a execução de um contrato. Nesse sentido, é bastante positivo que o rol de proteções proposto pela legislação brasileira seja substancialmente mais amplo do que o presente na regulação europeia, que inicialmente lhe serviu de inspiração." (Ibidem, p. 14).

[105] BURRELL, J. *How the Machine "Thinks:" Understanding Opacity in Machine Learning Algorithms*. Rochester, NY: Social Science Research Network, 2015. Disponível em: https://papers.ssrn.com/abstract=2660674. Acesso em: 27 maio. 2020.

ou serviços."[106] Essa definição é mais restrita que a definição do senso comum, presente nos dicionários, de que *automático* é aquilo que se move por conta própria, dispensando um operador. Outro sentido, figurado, remete à qualidade do que é involuntário e que ocorre independentemente da vontade.[107]

Esse duplo aspecto de ação sem aspectos volitivos, composto pela conjunção do sentido literal e do sentido figurado, é deveras ilustrativo das questões discutidas até agora. Pensemos no resultado produzido por um aplicativo que corresponda ao conceito de decisão anteriormente proposto, no caso, o Waze. Há várias rotas possíveis, das quais a aplicação seleciona as melhores segundo considerações realizadas pela plataforma e critérios estabelecidos pelo usuário. Há um processo automático, que se realiza sem a intervenção de um ente humano. Além disso, esse processo não se realiza pela vontade da aplicação. São critérios estabelecidos pelo usuário ou previamente desenvolvidos pela plataforma. Ou seja, o aspecto volitivo encontra-se fora da aplicação, no programador ou no usuário. Tal característica remonta à discussão apontada por Moor acerca da diferença entre a capacidade de decidir e a legitimidade ou competência para tal.

Sob o ponto de vista da criação de algoritmos, já argumentamos que se pode considerar que computadores têm a habilidade de decidir. Essa realidade, como demonstramos, inclui a capacidade de algumas aplicações desenvolverem os próprios critérios pelos quais realizarão tal escolha através do aprendizado de máquina. Contudo, o objetivo dessas aplicações, pelo qual a máquina buscará o aprendizado, depende do aspecto volitivo do programador. Neste sentido, embora possam operar de forma autônoma, continuam a depender da vontade humana em alguma medida para iniciarem suas atividades ou escolherem os objetivos.

Feita essa apresentação, podemos nos debruçar sobre o conceito de decisão automatizada presente na LGPD. O Art. 20 prevê que o "[...] titular dos dados tem direito a solicitar a revisão de decisões tomadas unicamente com base em tratamento automatizado de dados pessoais."[108] Podemos notar como o artigo, para o conceito

---

[106] ISA. WHAT IS AUTOMATION? Disponível em: https://www.isa.org/about-isa/what-is-automation. Acesso em: 14 jun. 2021.
[107] AUTOMÁTICO. *In*: MICHAELIS. São Paulo: Editora Melhoramentos, 2021. Disponível em: https://michaelis.uol.com.br/busca?id=bK57. Acesso em: 14 jun. 2021.
[108] BRASIL. *Lei nº 13.709, de 14 de agosto de 2018*. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF: Presidência da República, 14 ago. 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015- 2018/2018/lei/L13709.htm. Acesso em: 07. nov. 2020.

de decisão, apresenta o sentido estrito apontado por Moor[109], de uma decisão como uma escolha específica que permite o direito a uma revisão. Contudo, cumpre notar que inclui também os processos necessários para aquela conclusão, que, se forem realizados de forma unicamente automatizada de processamento de dados, fazem jus ao direito, portanto, em seu sentido amplo.

Pelo disposto no artigo, não há uma diferença para o caso de os critérios serem previamente programados por um ente humano ou criados de forma independente via aprendizado da máquina. Não há uma definição sobre o aspecto subjetivo da decisão, sobre quem desejou ou ordenou a decisão. O sentido para decisão está restrito à sua objetividade e ao processo pelo qual ocorreu. A legislação não especifica se a atividade de produção do conhecimento que baseou a decisão foi programada pelo agente de tratamento ou pelo algoritmo, dispondo apenas que se o tratamento foi realizado de forma automatizada, surge para o titular de dados um direito de revisão.

Nesse sentido, observamos que a redação do dispositivo, à semelhança do GDPR, assume uma postura pragmática, de que determinadas atividades de tratamento de dados, independentemente do seu aspecto volitivo, se realizada por meio de atividades automáticas, sejam computacionais ou não, serão consideradas como decisões automatizadas. Contudo, como todo processamento computacional de dados pessoais se realiza de forma automatizada, qualquer atividade computacional poderia ser considerada uma decisão. A inserção de um número de um cadastro num banco de dados, por exemplo, a partir da digitação, realiza-se por uma infinidade de operações aritméticas e rotinas de códigos computacionais cujo digitador não possui controle.

No entanto, depreende-se do art. 20 da LGPD que o conceito de decisão pressupõe que o resultado do processamento de dados seja contingente, ou seja, possua alternativas, de forma que não faria sentido falar em revisão de algo que permite apenas um resultado. No entanto, uma série de processamentos computacionais de dados pessoais permitem resultados alternativos. O envio de um terminal para um servidor de outra localidade ocorre pelo processamento automatizado, que divide a informação em pacotes e envia os *bits* por um caminho aleatório da rede. O caminho ocorre por uma série de decisões automatizadas orientadas por protocolos e poderia ser outro caso se o protocolo ou a configuração

[109] MOOR, J. Are There Decisions Computers Should Never Make. *Nature and System*, [*S. l.*], v. 1, 1985.

da rede fosse diversa.

A redação do artigo, então, utiliza uma perspectiva consequencialista e classifica a sua existência a partir de determinados efeitos. São decisões, para os efeitos do artigo, aquelas que afetam os interesses do titular. Neste sentido, o caminho que os *bits*, contendo informações pessoais, percorreram, embora automatizado e contingente, não possui, a princípio, efeitos para o titular dos dados. O artigo ainda inclui "[...] as decisões destinadas a definir o seu perfil pessoal, profissional, de consumo e de crédito ou os aspectos de sua personalidade"[110], como forma de apontar que tal efeito não precisa ser necessariamente material, podendo ser meramente potencial, como uma classificação sobre aspectos da personalidade.

Neste sentido, o conceito de decisão automatizada pode aqui ser construído como um resultado de um processamento de dados pessoais, sem a participação significativa de um operador humano, de forma computacional ou não, que produza ou possa produzir efeitos no indivíduo e cujo processamento de dados em questão permita outros resultados possíveis.

Por fim, como demonstraremos ao longo deste trabalho, um algoritmo que toma determinada decisão, ainda que de maneira autônoma, não o faz desvinculado de toda uma cadeia de processamentos de dados e relações jurídicas ou comerciais. Em nossa acepção ampla de decisão, que inclui os procedimentos necessários à escolha, é preciso destacar que devem ser compreendidas como um processo complexo, do qual o resultado do processamento automático constitui apenas um momento. Portanto, é importante considerar o processamento de dados em toda a cadeia que engloba um sistema sociotécnico.[111] Nesse sentido, o art. 20 da LGPD, assim como o art. 22 do GDPR, causou alguma preocupação a respeito da possível interpretação de que a menor participação humana em uma decisão seria suficiente para afastar o direito à revisão[112], pelo que entendemos necessário partir de uma abordagem que considere decisão automatizada não aquela decisão privada de toda e qualquer

---

[110] BRASIL*. Lei nº 13.709, de 14 de agosto de 2018*. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF: Presidência da República, 14 ago. 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015- 2018/2018/lei/L13709.htm. Acesso em: 07. nov. 2020.

[111] MITTELSTADT, B. D. *et al.* The ethics of algorithms: Mapping the debate. *Big Data & Society*, [*S. l.*], v. 3, n. 2, 2016. Disponível em: https://doi.org/10.1177/2053951716679679. Acesso em: 20 maio 2020. p. 205395171667967.

[112] SOUZA, C. A.; PERRONE, C.; MAGRANI, E. O direito à explicação entre a experiência europeia e a sua positivação na LGPD. *In*: MENDES, L. S.; DONEDA, D.; SARLET, I. W.; RODRIGUES JÚNIOR, O. L.; BIONI, B. (Orgs.). *Tratado de Proteção de Dados Pessoais*. Rio de Janeiro: Forense, 2021. p. 243-270.

participação humana, mas aquelas decisões tomadas com base no tratamento automatizado de dados não sujeitas a uma intervenção ou participação humana significativa.[113]

A definição apresenta, sem dúvida, uma previsão importante no contexto de novas tecnologias emergentes, baseadas no uso de dados, que automatizam uma série de atividades antes realizadas por seres humanos. Embora existisse precedentes na legislação brasileira para a explicabilidade, como as previsões do Código de Defesa do Consumidor e da Lei de Cadastro Positivo[114], das quais trataremos no Capítulo 3, a previsão da LGPD é explícita e estende a proteção para além das relações consumeristas e menciona o direito de revisão. No entanto, autores apontam como tal previsão de revisão de decisões automatizadas pode não ser suficiente para mitigar os riscos a direitos fundamentais que emergem com essas tecnologias.[115]

2.1.2 O que é uma explicação no contexto das decisões automatizadas?

Uma vez discutidos os contornos e as possibilidades de um direito à explicação, bem como o que viriam a ser, efetivamente, decisões automatizadas, o presente tópico passará a discutir, de forma pormenorizada, o que seria uma explicação no contexto de decisões automatizadas. Para tanto, serão analisadas, em um primeiro momento, as diferenças entre o direito à explicação em relação às obrigações de transparência. Em seguida, passaremos a discutir a ideia de compreensão e quais seriam os tipos de explicação possíveis, pontuando seus objetivos e suas formas. Por fim, analisaremos quais os requisitos necessários para que uma explicação seja considerada efetiva, bem como quais seriam seus objetivos gerais e imediatos.

Inicialmente, cabe pontuar que o direito à informação é um pressuposto essencial de todo processo ao qual pessoas são submetidas. Até por isso que o direito à informação é um dos corolários do devido processo legal, que visa, no mínimo, garantir proporcionalidade e razoabilidade nas eventuais limitações de direitos. Desta

---

[113] Sobre a noção de participação humana significativa, cf. Capítulo 3.

[114] MONTEIRO, R. L. *Existe um direito à explicação na Lei Geral de Proteção de Dados do Brasil?* Rio de Janeiro: Instituto Igarapé, dez. 2018. Artigo Estratégico nº 39. Disponível em: https://igarape.org.br/wp-content/uploads/2018/12/Existe-um-direito-a-explicacao-na-Lei-Geral-de-Protecao-de-Dados-no-Brasil.pdf. Acesso em: 21 jun. 2021.

[115] WACHTER, S.; MITTELSTADT, B. A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI. *LawArXiv*, 12 out. 2018. Disponível em: https://doi.org/10.31228/osf.io/mu2kf. Acesso em: 27 maio 2020.

forma, o direito à informação é um passo necessário rumo à garantia de um direito à explicação em face de decisões automatizadas. No entanto, antes é necessário se debruçar sobre o papel que a transparência, que orienta o direito à informação, ocupa nos processos de *accountability*. Sem a devida abordagem sistêmica, a transparência pode não garantir os efeitos prometidos.

Em 2016, Ananny e Crawford escreveram o *paper* “Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability”[116], cujo objetivo é fazer uma crítica fundamental à centralidade da transparência como elemento de *accountability.* O texto parte da ideia de que modelos utilizados para entender e responsabilizar (“*hold accountable*”) há muito têm se apoiado na ideia de transparência, como se a possibilidade de ver um sistema significasse necessariamente entender seu funcionamento e governá-lo. A pergunta que norteia o artigo é: “caixas pretas” podem ser abertas, e, caso possam, isso é suficiente?

Segundo os autores, a demanda por maior transparência para sistemas algorítmicos se fortalece conforme estes sistemas são mais utilizados pelo poder público. Historicamente, da filosofia ao ativismo, a transparência tem sido mobilizada como uma forma de saber a “verdade” de um sistema, a partir da ideia implícita de que ver algo permite que se mude algo. A sugestão dos autores é que, ao invés do foco em ver sistemas *por dentro,* se busque *accountability* por meio de um olhar transversal deles — enxergando-os como sistemas sociotécnicos que não *contêm* complexidade, mas *promovem* complexidade.

Podemos identificar uma lógica na sustentação da visão tradicional sobre a transparência. Nessa forma de ver, a observação produz um *insight,* que por sua vez cria conhecimento necessário para governar e garantir a *accountability* de sistemas. Tem por base uma premissa epistemológica de que a verdade equivale à correspondência com um fato, que o objetivo da investigação científica é sempre *descobrir* uma verdade encoberta. Tal ideia predominante relaciona-se à tradição iluminista na fundação das ciências sociais e sua associação com métodos aplicados nas ciências naturais.

A transparência, portanto, incluiria uma “promessa de controle” a partir da observação e conhecimento de uma determinada realidade. Esta premissa,

[116] ANANNY, M.; CRAWFORD, K. Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, v. 20, n. 3, p. 973–989, 2016.

entretanto, se apoia na ideia de que a informação “descoberta” seja facilmente acessível e discernível por uma determinada audiência. Além disso pressupõe que os indivíduos ajam para mudar e garantir a promessa de futuro que esta abertura sugere ser possível. Os autores discordam de que isso corresponda à realidade.

Eles sugerem que utilizemos três metáforas para a transparência identificadas por estudiosos do conhecimento: “valor público incorporado pela sociedade para combater a corrupção”, sinônimo de “processo decisório aberto por parte de governos e organizações não-lucrativas” e “ferramenta complexa de boa governança”[117]. O entusiasmo em torno deste valor, segundo os autores, penetrou profundamente no campo da pesquisa sobre cultura de participação *online*, a partir de “suposições não examinadas” sobre os benefícios da transparência e a equivalência entre a capacidade de enxergar um sistema por dentro e a capacidade de governá-lo. A mesma lógica dá substância ao conceito de “transparência algorítmica”, entendida por muitos estudiosos como uma forma adequada de combate à discriminação e às incertezas que estes sistemas carregam, de maneira a possibilitar uma governança efetiva sobre eles.

A primeira tese apresentada pelos autores que desafia a premissa epistemológica da transparência é a que considera a verdade como significados construídos por meio de relações e não de revelações. O raciocínio que embasa essa tese é o de que, ainda que descrições formais de fenômenos observáveis sejam valiosas, ideias “[...] tornam-se verdade apenas na medida em que se relacionam de forma satisfatória com outras partes da experiência humana.”[118] A ideia de “*accountability* por meio da visibilidade”, por outro lado, se apoia na noção de verdade pela aparência, e não pela correspondência. A ilusão da transparência, nesse sentido, é a promessa de uma *accountability* que a transparência não pode de fato proporcionar.

A partir desta introdução, os autores elaboram um rol de dez principais limitações da noção de transparência, conforme desenvolvida na literatura tradicional:

1. **A transparência pode ser desconectada do poder**. Se a visibilidade sobre

[117] BALL, C. What Is Transparency? *Public Integrity*, [*S. l.*], v. 11, n. 4, p. 293–308, 2009. Disponível em: https://doi.org/10.2753/PIN1099-9922110400. Acesso em: 21 jun. 2021.
[118] JAMES, W. What pragmatism means. *In:* MENAND, L. *Pragmatism: A Reader*. New York: Random House, 1997. p. 93–111.

uma questão, como a corrupção, por exemplo, não tem alguma consequência prática, ela pode perder o sentido e até criar um clima de ceticismo. A ideia de que a transparência produz mudança depende da premissa de que quem detém o poder sobre essa mudança é de fato vulnerável à transparência;

2. **A transparência pode ser prejudicial.** Se não houver clareza sobre os porquês da transparência, ela pode trazer danos à privacidade dos indivíduos e, com isso, inibir um diálogo honesto;
3. **A transparência pode causar opacidade.** Intencionalmente ou não, a liberação de grandes volumes de informação pode efetivamente esconder os pedaços dessa informação que são mais relevantes/úteis para a garantia de *accountability,* no meio de uma "pilha de informação inútil";
4. **A transparência pode criar falsos binários.** Sem entendimentos sutis sobre o tipo de responsabilidade que a visibilidade é projetada para criar, as chamadas à transparência podem ser lidas como falsas escolhas entre o sigilo total e a total abertura;
5. **A transparência pode invocar modelos neoliberais de agência.** A premissa da transparência é de um "mercado iluminado da informação", isto é, a noção de que dar informação às pessoas significa que elas terão condições de empregá-las para tomar decisões melhores, que, por sua vez, levarão a resultados sociais desejáveis;
6. **A transparência não necessariamente garante confiança.** Não há evidências empíricas suficientemente sólidas para demonstrar que maior transparência necessariamente gera maior confiança, pois isso varia conforme os diferentes atores e contextos de interação social;
7. **A transparência pode envolver limitações corporativistas.** Comumente a transparência é inviabilizada pois profissionais tendem a proteger a exclusividade do seu trabalho e *expertise;*
8. **A transparência pode privilegiar o *ver* em detrimento do *entender.*** Partindo de conceitos de teorias educacionais, como as de Piaget e Vygotsky, que preconizam não apenas a demonstração da existência de sistemas para crianças, mas seu envolvimento direto nos processos que caracterizam esses sistemas, os autores alegam que a visão engessada de transparência resulta em um menor, e não maior entendimento, sobre os sistemas que busca revelar;

9. **A transparência tem limitações técnicas.** A escala e velocidade do *design* de um sistema pode torná-lo opaco e até ininteligível, inclusive para seus próprios criadores, como os modelos conhecidos como *deep learning*;
10. **A transparência tem limitações temporais.** Objetos e sistemas mudam ao longo do tempo, o que ganha especial relevância quando se fala de *softwares* e algoritmos e, mais ainda, diante de sistemas altamente adaptativos. Dessa forma, ainda que o código-fonte e dados de treinamento de um determinado algoritmo fossem transparentes, ainda assim isso representaria apenas uma "fotografia" de sua funcionalidade. Além disso, os autores ressaltam que instrumentos e representações do conhecimento não são separados das práticas e culturas daqueles que os observam/estudam.

Os autores sugerem que a exigência de abertura da "caixa preta" é uma metáfora ruim e insuficiente para abordar a complexidade dos sistemas algorítmicos contemporâneos. Um sistema precisa ser *compreendido* para ser governado — e essa compreensão deve ser ampla, incluindo a maior quantidade de entendimentos possível.

Em contextos digitais, transparência não é uma questão apenas de revelar informações, mas um processo dinâmico de configuração e reconfiguração de plataformas, algoritmos e sistemas de *machine learning* que gerenciam visibilidade. Tais "dispositivos de divulgação"[119] não são exclusivamente humanos nem inteiramente computacionais, mas redes de agentes humanos e não-humanos que criam "visibilidades e possibilidades particulares de observação".

Assim, garantir a *accountability* de um conjunto é mais do que olhar para os seus componentes individualmente, ou mesmo para um retrato do todo em um determinado momento, mas requer que se *entenda* o funcionamento do sistema e como seus elementos se *relacionam* entre si e com o ambiente. Isso leva a uma reconfiguração da pergunta inicial, que deveria ser: que tipos de reivindicações podem ser feitas para *entender* uma rede de atores e como esse entendimento está relacionado com simplesmente *ver* uma rede de atores?

Um ponto positivo no debate sobre as limitações da transparência é que, para além de uma mera discussão, o conhecimento destas limitações deve sugerir o

[119] HANSEN, H. K.; FKYVERBOM, M. The politics of transparency and the calibration of knowledge in the digital age. *Organization*, v. 22, n. 6, p. 872–889, 2015.

próprio caminho que o modelo de *accountability* pode seguir. Se a transparência é inefetiva porque o poder é imune à visibilidade, então o modelo deve focar justamente na assimetria de poder que é revelada por esta maior transparência. A título de fechamento, os autores propõem:

> Se um sistema é tão complexo que mesmo aqueles com uma visão total sobre ele são incapazes de descrever suas falhas e sucessos, então os modelos de *accountability* devem focar em se o sistema é suficientemente compreendido ou compreensível para permitir sua implementação em diferentes ambientes, se mais tempo de desenvolvimento é necessário, ou mesmo se o sistema *deve* ser implementado.[120]

A partir das considerações elaboradas por Ananny e Crawford, tornam-se visíveis as inúmeras limitações do direito à transparência frente ao contexto de decisões automatizadas. Mesmo quando as obrigações de transparência passiva e ativa garantam aos titulares dos dados o direito de acesso à integralidade dos dados pessoais que uma organização, pública ou privada, detém sobre ele, o mero fornecimento desses dados provavelmente não permitirá ao titular ter um conhecimento efetivo sobre a forma como são tratados e para as finalidades para as quais são tratados. Isso porque estas situações encontram limitações na cognição do próprio titular para compreender as informações que lhe são repassadas, principalmente as relativas à lógica subjacente ao processamento dos dados, conforme discutiremos no Capítulo 4 deste trabalho. Portanto, o direito de acesso, em si, mesmo atrelado às obrigações de transparência, pode não ser suficiente para garantir o efetivo controle sobre o fluxo de dados e a autodeterminação informativa, de forma que se deve reconhecer um verdadeiro direito à explicação.

Uma vez esclarecidas as razões pelas quais não se deve confundir o direito à explicação com as obrigações de transparência, passaremos a discorrer, de forma resumida, sobre a noção de *compreensão* no contexto do direito à explicação em face de decisões automatizadas, discussão que será abordada com mais detalhes no item 4.2 deste trabalho.

A ideia de garantir ao titular a efetiva compreensão acerca do funcionamento, da lógica subjacente ao sistema algorítmico e dos efeitos de uma decisão automatizada que venham a impactar significativamente sua esfera de direitos exige

---

[120] ANANNY, M.; CRAWFORD, K. Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, v. 20, n. 3, p. 973–989, 2016.

que se vá além da mera disponibilização de informações, da mera abertura à transparência e do mero atendimento a uma requisição de acesso aos dados. A garantia de um direito à explicação exige que o agente responsável não apenas apresente, como também articule as informações que se façam necessárias para a efetiva compreensão do funcionamento de determinado sistema automatizado, o que inclui, dentre outras, as informações relativas aos fundamentos da decisão, aos *inputs* fornecidos e aos critérios que conduziram à decisão automatizada.

Mais recentemente, em 2019, Bohlender e Kohl apresentaram o *paper "*Towards a Characterization of Explainable Systems"[121], cuja premissa é que, a despeito do aumento da complexidade e autonomia de sistemas movidos por *softwares* e do direcionamento dos esforços de pesquisa atuais para o desenho de *sistemas explicáveis,* não se tem, na literatura, um entendimento comum sobre o que é de fato necessário para tornar um sistema explicável. O artigo pretende, então, reunir os avanços recentes desta discussão a fim de consolidar uma terminologia comum.

Introdutoriamente, os autores exploram as mais diversas razões para se buscar sistemas explicáveis. Eles permitem, por exemplo, que engenheiros localizem e consertem *bugs* e que haja cooperação entre empresas e usuários, na medida em que deixar os consumidores de um sistema "no escuro" prejudica a confiança e, além disso, pode levar a problemas e erros operacionais[122]. No caso de sistemas automatizados de *machine learning,* entretanto, essa relação é muito complexa e, muitas vezes, não é claro sequer *qual tipo* de explicação seria considerado apropriado. É evidente que, nessa área, há explicações, mas elas são técnicas e inacessíveis para pessoas leigas.

Adiante, o artigo argumenta que o ímpeto por obter uma explicação decorre da premissa simples de que falta entendimento sobre algum aspecto de um sistema, o que leva à busca por uma determinada representação de uma informação que seja considerada satisfatória. Dessa forma, a explicação pode ser definida como a representação de uma informação, que explica um aspecto de um sistema — o *explanandum* — para um indivíduo ou um grupo, na medida em que o processamento

[121] BOHLENDER, D.; KÖHL, M. A. Towards a Characterization of Explainable Systems. *ArXiv [cs]*, [*S. l.*], 2019. Disponível em: http://arxiv.org/abs/1902.03096. Acesso em: 6 out. 2020.
[122] BOHLENDER, D.; KÖHL, M. A. Towards a Characterization of Explainable Systems. *ArXiv* [cs], [*S. l.*], 2019. Disponível em: http://arxiv.org/abs/1902.03096. Acesso em: 6 out. 2020.

desta representação por qualquer membro representativo deste grupo faz com que este membro de fato entenda o aspecto do sistema em questão.

Aqui, faz-se uma distinção entre a informação e a representação da informação, pois uma mesma informação pode ter representações variadas. A informação é factual, a representação pode ser distorcida e a explicação, por sua vez, é aquela representação que carrega informações relevantes e necessárias para que o indivíduo possa compreender determinado *explanandum.* Essa noção de informação também independe do agente envolvido, enquanto a representação de uma informação se altera de acordo com quem a projeta e produz.

Dentro da categorização proposta, o *processamento* da representação de uma informação é relevante, na medida em que a explicação só se o agente representativo de um determinado grupo é capaz de processar a representação, enquanto a representação faz algum sentido para esse agente. Tal processamento pode ser influenciado pelo contexto, pelo tempo disponível para que o agente possa apreender a representação, dentre outros fatores. Por outro lado, os autores salientam que o mero processamento de uma representação (seja por via cognitiva ou computacional) não a torna uma explicação, pois o processamento deve fazer com que o agente entenda de fato o aspecto da informação relevante.

E o que significa entender? O texto defende que compreensão é um conceito mais tangível do que explicação por conta do extenso acúmulo acadêmico em áreas como a psicologia e a ciência comportamental. Partindo da literatura nessas áreas, a autoridade de determinar o que é “entendível” pertence ao grupo visado e seus membros, isto é, aos indivíduos a quem a explicação em tese se destina.

Outro ponto que Kohl e Bohlender fazem questão de sublinhar é que não é suficiente que um agente *pense* que entendeu determinada representação, este entendimento deve ser genuíno. Uma ideia de como mensurar este parâmetro complexo é medir os efeitos de uma explicação. Como exemplo, remete à hipótese do engenheiro em busca de erros em um sistema e aponta que uma explicação seria considerada suficiente caso tal engenheiro fosse efetivamente capaz de localizar os referidos erros depois de processar uma explicação a respeito.

A necessidade de explicação, conforme previamente mencionado, decorre do fato de que algo não é entendido, mas o texto aponta que é necessário um avanço substancial nos estudos sobre *como descrever precisamente uma falta de compreensão*. Importante destacar também que, para o mesmo *explanandum*, pode

haver diferentes explicações em diferentes dimensões e que a ideia de criar um vocabulário comum sobre o que significa entender e sobre como descrever a falta de compreensão acerca de determinado aspecto de um sistema também permite que haja comparações melhores entre diferentes explicações em um determinado campo.

Adiante, o texto aborda a questão do grupo visado, uma vez que, dentro de um grupo utilizado com fim de generalização pode haver agentes específicos que não tenham condições de entender determinadas representações. Por esse motivo, a caracterização proposta propõe como parâmetro um "agente representativo" de um determinado grupo, isto é, agentes equipados com o *background* e conhecimento necessários para aquele grupo.

Para prover acesso a explicações, é necessário que elas sejam produzidas por algo ou alguém. O meio para produzir uma explicação pode ser o mesmo sistema (sistemas auto-explicáveis), outro sistema, um ser humano, etc. Esse ponto é relevante, para os autores, pois a mera existência teórica de uma explicação possível não significa que um sistema seja de fato explicável, é preciso ter uma materialização. Isso significa que não só a explicação é relativa (a um sistema, um aspecto, um grupo) como um sistema também não é explicável *em si,* mas em relação a certos aspectos e um grupo visado.[123]

Quando não há meios apropriados capazes de produzir uma explicação de um aspecto de um sistema, então este aspecto é considerado inexplicável. Os autores reconhecem que nem tudo é, de fato, explicável — nem mesmo em teoria — e mesmo que o seja em teoria, por vezes o esforço prático demandado para desenhar um meio apropriado para produzir essa explicação pode ser excessivo.

A última seção do texto aborda as disciplinas do conhecimento que podem ser envolvidas nas várias fases do processo e das distintas caracterizações descritas. Sobre representações, por exemplo, a ciência cognitiva e a psicologia, sobre interpretabilidade e como desenhar *softwares* que sejam explicáveis a ciência da computação e a engenharia de *softwares*. A inteligência artificial[124], por seu turno, tem todo um desafio próprio de explicar os sistemas desenvolvidos com base na sua lógica, cada vez mais complexa em razão do avanço dos sistemas de *deep learning*.

---

[123] BOHLENDER, D.; KÖHL, M. A. Towards a Characterization of Explainable Systems. *ArXiv* [cs], [*S. l.*], 2019. Disponível em: http://arxiv.org/abs/1902.03096. Acesso em: 6 out. 2020.

[124] Em relação a estas últimas tecnologias, já há um campo profícuo de pesquisa, denominado "*Explainable Artificial Intelligence*" (XAI).

Mas o que viria a ser, de fato, uma explicação no contexto de decisões automatizadas e quais seriam seus objetivos e requisitos? Conforme apontam Brkan e Bonnet, por vezes, é difícil traçar uma definição precisa do que seria uma explicação em abstrato, sem levar em consideração aspectos contextuais como o modelo de decisão automatizada em questão, os tipos de dados utilizados, a compreensão almejada com aquela explicação ou o tipo de perguntas que fazemos a respeito da decisão. Essa dificuldade de definição e o caráter multifacetado das explicações levaram alguns autores a formularem classificações para diferentes tipos de informações, que encerram em si diferentes sentidos sobre o que viria a ser uma explicação e diferentes métodos de como alcançá-la.[125]

Dentre essas classificações, tendo em vista o objeto da explicação, a literatura costuma apontar para dois tipos de explicações: (i) explicações globais, voltadas a explicar os mecanismos de funcionamento de um sistema automatizado como um todo, e explicações locais, destinadas a fornecer informações sobre os fundamentos e os aspectos que levaram à formação de uma decisão específica.[126] Nesse sentido, explicações locais teriam como objetivo explicitar os principais aspectos de uma decisão em particular, traçando correlações entre os *inputs* fornecidos e os *outputs* gerados. Informações globais, por sua vez, objetivam esclarecer como determinado sistema automatizado opera internamente, devendo conter informações relacionadas à lógica de funcionamento do modelo como um todo. Sobretudo quanto às explicações que recaem sobre o modo de funcionamento dos sistemas, há inúmeras limitações de natureza legal (como direitos de propriedade industrial visando proteger o segredo de negócio em torno do código fonte do modelo algorítmico) e técnicas (como a própria complexidade e imprevisibilidade de sistemas como *machine learning*) que acabam por limitar a extensão dessas explicações.

Essas e outras limitações também foram apontadas por Wachter *et al.* em seu artigo “Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR”[127]. A despeito destas limitações, os autores sustentam que o valor da explicabilidade de sistemas de decisões automatizados persiste, mas que

---

[125] BRKAN, M.; BONNET, G. Legal and Technical Feasibility of the GDPR’s Quest for Explanation of Algorithmic Decisions: of Black Boxes, White Boxes and Fata Morganas. *European Journal of Risk Regulation*, v. 11, n. 1, 2020. p. 25-26.
[126] Ibidem, p. 32.
[127] WACHTER, S.; MITTELSTADT, B.; RUSSELL, C. Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR. *SSRN Electronic Journal*, v. 31, n. 2, 2017. Disponível em: https://www.ssrn.com/abstract=3063289. Acesso em: 27 maio. 2020.

um ponto que consideram central tem sido negligenciado neste debate: que uma explicação de decisões automatizadas, e um direito geral de explicação, não dependem, necessariamente, de que o público *entenda* como um algoritmo funciona.

Em outras palavras, o objetivo do artigo é demonstrar que, *a priori*, explicações podem ser fornecidas sem que se abra a "caixa preta", a partir da proposta de três objetivos principais que uma explicação deve cumprir: (i) informar e auxiliar o titular a compreender por que uma determinada decisão foi atingida; (ii) fornecer a base para a contestação de uma decisão; e (iii) compreender o que pode/deve ser alterado para que um resultado diferente seja obtido no futuro. A solução apresentada pelos autores para que todos os três objetivos sejam cumpridos consiste em fornecer "explicações contrafactuais", isto é, um raciocínio construído a partir de orações condicionais em que uma delas é falsa.

Diferentemente da lógica que permeia a literatura sobre a explicação no contexto de sistemas automatizados, essa proposta baseia-se em elementos externos que conduzem a uma decisão e não sua lógica interna (se fator x fosse diferente, então determinada classificação de um indivíduo seria y). Neste sentido, os autores defendem que a GDPR não cria um direito que requeira o destrinchamento dos sistemas de decisões automatizadas, de forma que a proposta de abordagem deste *paper* se amoldaria às exigências dos seus dispositivos.

O art. 12 (7)[128] do Regulamento, segundo os autores, corrobora esta tese, na medida em que esclarece que o objetivo dos arts. 13 e 14 é "[...] de forma visível, inteligível e legível oferecer um panorama significativo do tratamento pretendido."[129] O art. 12(1), por sua vez, estabelece que toda comunicação e informação destinada ao titular de dados deve ser fornecida de maneira "concisa, transparente, inteligível e

---

[128] "Article 12 EU GDPR – 'Transparent information, communication and modalities for the exercise of the rights of the data subject'.
[...] 7. The information to be provided to data subjects pursuant to Articles 13 and 14 may be provided in combination with standardised icons in order to give in an easily visible, intelligible and clearly legible manner a meaningful overview of the intended processing. Where the icons are presented electronically, they shall be machine-readable." (UNIÃO EUROPEIA. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance)*. Official Journal of the European Union, 4 maio 2016. Disponível em; https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng. Acesso em: 27 ago. 2021).

[129] UNIÃO EUROPEIA. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance)*. Official Journal of the European Union, 4 maio 2016. Disponível em; https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng. Acesso em: 27 ago. 2021.

facilmente acessível", o que sugere que uma abordagem complexa e baseada em "juridiquês" ou explicações altamente técnicas seria considerada inadequada.

As explicações contrafactuais, neste sentido, serviriam ao objetivo de explicitar a função, o impacto ou o valor de diferentes tipos de informação em uma dada decisão automatizada. Conforme assinalam Brkan e Bonnet:

> Furthermore, local explanations are suitable in terms of GDPR requirements. Local explanations probe the system to determine the correlations between input and output as well as to extract the main factors of the decision. While this approach individualises the explanations, it can be quite time consuming if the system is probed for every decision of every individual. *Counter-factual faithfulness functions in a similar fashion, with the difference being that it evaluates how a change in a particular factor influences the decision. It may be used, for instance, to assess the fairness of a decision, based on how each factor within a given specific input influences the decision*.[130] (grifo nosso)

Ainda sobre os requisitos e objetivos de uma explicação efetiva, Selbst e Powles[131], ao discorrerem sobre a interpretação do termo *meaningful information* no texto da GDPR, defendem que, ainda que a interpretação do que seria "informação significativa" dependa de uma construção ao longo do tempo, algumas observações já podem ser feitas em relação à sua aplicabilidade. A primeira delas é que este termo se relaciona aos titulares de dados, isto é, que a informação deve ser significativa/relevante *para o titular,* alguém que se supõe não ser um *expert* em tecnologia e computação. Em segundo lugar, os autores sustentam que essa análise deve ser funcional, isto é, deve observar o que a explicação em questão agrega à capacidade do titular de se manifestar contra uma decisão, o que seria, para eles, uma abordagem mais relevante do que a que confere um valor intrínseco a este direito, de difícil mensuração.

Não constitui objetivo deste trabalho discorrer extensivamente sobre todas as modalidades de explicação, seus objetivos e requisitos, mas tão somente esclarecer que há diferentes formas de se chegar a uma explicação, por meio de diferentes metodologias, e que cada opção se encontra relacionada a objetivos e resultados

---

[130] BRKAN, M.; BONNET, G. Legal and Technical Feasibility of the GDPR's Quest for Explanation of Algorithmic Decisions: of Black Boxes, White Boxes and Fata Morganas. *European Journal of Risk Regulation*, v. 11, n. 1, 2020. p. 49.

[131] SELBST, A. D.; POWLES, J. Meaningful Information and the Right to Explanation. *International Data Privacy Law*, v. 7, n. 4, p. 233-242, 2017. Disponível em: https://papers.ssrn.com/abstract=3039125. Acesso em: 27 maio 2020. p. 7-9.

específicos. Brkan e Bonnet apresentam um quadro geral das diferentes modalidades de explicação, explicitando seus objetivos e sua viabilidade técnica e jurídica:

QUADRO 1 – Viabilidade técnica e jurídica de modalidades de explicação algorítmica

TIPO DE EXPLICAÇÃO	SIGNIFICADO DE EXPLICAÇÃO	TEMPO	MODELO MAIS ADEQUADO PARA EXPLICAR	MÉTODO TÉCNICO MAIS ADEQUADO	GRAU DE VIABILIDADE TÉCNICA	GRAU DE VIABILIDADE PARA SUPERAR OBSTÁCULOS LEGAIS	GRAU DE CORRESPONDÊNCIA COM AS EXPLICAÇÕES DO GDPR
Checagem de propriedade	Provar que uma decisão irá sempre ou nunca ocorrer em determinadas situações	Ex ante	Não supervisionado	Métodos de *white-box* externa	Moderado	Difícil	Baixa
Traços interpretáveis	Traduzir traços da execução para linguagem natural	Ex post	Supervisionado	Métodos reflexivos *built on top*	Difícil	Moderado	Alto
Explicações locais	Identificar os principais fatores envolvidos	Ex ante / Ex post	Não supervisionado	Métodos de Black-box Externa	Fácil	Fácil	Alto
Fidelidade contrafactual	Avaliar a influencia de diferentes fatores	Ex ante / Ex post	Não supervisionado	Métodos de Black-box Externa	Fácil	Fácil	Alto
Descompilação da decisão	Aproximação da lógica envolvida	Ex ante / Ex post	Não supervisionado	Métodos de Black-box Externa	Fácil	Fácil	Alto
Fornecer argumentos	Prover argumentos favoráveis e contrários para uma decisão	Ex post	Supervisionado / Não supervisionado	Métodos reflexivos *built within*	Difícil	Moderado	Médio

Fonte: Brkan; Bonnet, 2020, p.47.

No presente tópico, buscamos apresentar as diferentes modalidades ou formas de operacionalizar uma explicação, destacando suas vantagens, desvantagens e limitações. A partir da discussão apresentada ao longo desta seção e a partir do quadro geral traçado por Brkan e Bonnet acima exposto, é possível constatar que não há uma única forma de explicação, mas sim diferentes métodos e técnicas que podem ser empregados para explicar o funcionamento de determinado sistema automatizado. A escolha de determinada forma de explicação deverá ser realizada com base numa análise contextual, considerando-se as características do sistema que se busca compreender e explicar, os objetivos que se pretende alcançar com determinada explicação, o *timing* da explicação (*se ex ante*, *ex post* ou ambas), dentre outros fatores.

Uma vez discutidas as diferentes formas de explicação, passaremos a debater especificamente como é possível fundamentar a existência de um direito à explicação, seja mediante previsões normativas expressas, seja mediante a interpretação sistemática de princípios e direitos positivados, seja a partir de elementos e

fundamentos metajurídicos.

2.2 COMO FUNDAMENTAR A EXISTÊNCIA DE UM DIREITO À EXPLICAÇÃO?

Como visto na seção 2.1, apesar da importância e essencialidade do direito à explicação, existem dúvidas sobre a sua real existência na extensão necessária para atingir os objetivos almejados. Neste ponto, existiria uma diferença entre o "*right to notice*", que visa informar o titular dos dados sobre a forma como seus dados pessoais serão tratados, e o "*right to explanation*", que incluiria o primeiro direito, mas com um escopo consideravelmente maior, abarcando explicação, de uma maneira que leve em consideração as limitações cognitivas do titular, sobre a lógica por trás do tratamento dos dados pessoais, e, ainda, os possíveis impactos que tal tratamento pode ter nos seus direitos e nos de terceiros.

A fim de fundamentar a existência do direito à explicação sobre decisões automatizadas no ordenamento jurídico brasileiro, é necessário, primeiro, circunscrever a pesquisa a um modelo teórico robusto sobre como indivíduos adquirem direitos subjetivos. Há alguns caminhos que podem ser trilhados nesse sentido: i) direitos subjetivos "nascem" a partir de explícita previsão legal; ii) direitos subjetivos/princípios podem ser extraídos de um conjunto de normativas que disciplinam direitos mais amplos (como a privacidade ou o direito à informação, neste caso), a partir de uma interpretação teleológica e sistemática; e iii) direitos subjetivos podem ser derivados de valores morais, de acordo com determinadas concepções da ordem da Ética, Filosofia Política e Filosofia do Direito e de acordo com os fundamentos de legitimidade da própria ordem jurídica: se a autoridade normativa do direito é derivada, por exemplo, do tratamento de todos os cidadãos com dignidade, respeito e consideração, as leis devem ser interpretadas de modo a garantir esses valores nos casos concretos aos quais se aplicam.

Essas três formas pelas quais é possível reconhecer a existência de direitos subjetivos podem ser mais bem compreendidas a partir da Teoria do Ordenamento Jurídico formulada por Norberto Bobbio. Ao examinar o ordenamento jurídico, abstratamente tomado enquanto um sistema de normas, o autor destaca a existência de três aspectos que lhes são essenciais: a unidade, a coerência e a completude. O problema que aqui buscamos enfrentar pode ser entendido a partir da categoria da completude, entendida como:

> [...] a propriedade pela qual um ordenamento jurídico tem uma norma para regular qualquer caso. Uma vez que a falta de uma norma se chama geralmente "lacuna" (num dos sentidos do termo "lacuna"), "completude" significa "falta de lacunas". Em outras palavras, um ordenamento é completo quando o juiz pode encontrar nele uma norma para regular qualquer caso que se lhe apresente, ou melhor, não há caso que não possa ser regulado com uma norma tirada do sistema.[132]

Em última análise — argumenta o autor — um ordenamento jurídico, se *estaticamente* considerado, pode não ser *completo*, mas se *dinamicamente* considerado, pode ser tomado como *completável*. Valendo-se da doutrina italiana formulada por Carnelutti, Bobbio afirma haver ao menos duas formas ou dois métodos de complementação de um ordenamento jurídico: a heterointegração e a auto-integração. A primeira técnica consiste em recorrer a ordenamentos diversos ou a fontes diversas daquela que se entende como dominante (a lei).[133] Neste sentido, ao menos durante o período no qual as correntes jusnaturalistas foram predominantes, era comum que os juízes, no caso de lacuna do ordenamento, recorressem ao direito natural na tentativa de superação das lacunas existentes no direito positivo. Ainda no método de heterointegração, Bobbio menciona a possibilidade de recurso a outros ordenamentos positivos, aos costumes, aos juízos de equidade, ou seja, ao poder criativo do juiz, e à doutrina como elementos úteis à tarefa de integração.[134] "O segundo método" – afirma – "consiste na integração cumprida através do mesmo ordenamento, no âmbito da mesma fonte dominante, sem recorrência a outros ordenamentos e com o mínimo recurso a fontes diversas da dominante."[135] O método da auto-integração, assim, apoia-se no recurso a dois elementos[136]: a analogia, entendida como "[...] o procedimento pelo qual se atribui a um caso não-regulamentado a mesma disciplina que a um caso regulamentado semelhante"[137] e os princípios gerais do direito, entendidos não como princípios gerais extraídos de um direito natural, mas os princípios gerais fundantes de um ordenamento positivado. Bobbio apresenta uma interessante classificação dos princípios gerais em (i)

---

[132] BOBBIO, N. *Teoria do ordenamento jurídico*. Apresentação: Tércio Sampaio Ferraz Júnior; Tradução: Maria Celeste C. J. Santos; Rev. téc.: Cláudio de Cicco. 6. ed. Brasília: Editora Universidade de Brasília, 1995. p. 115.
[133] Ibidem, p. 146.
[134] Ibidem, 147-150.
[135] Ibidem, p. 147.
[136] Ibidem, p. 150.
[137] Ibidem, p. 151.

princípios gerais expressos e (ii) princípios gerais não expressos,

> [...] ou seja, aqueles que se podem tirar por abstração de normas específicas ou pelo menos não muito gerais: são princípios ou normas generalíssimas, formuladas pelo intérprete, que busca colher, comparando normas aparentemente diversas entre si, aquilo a que comumente se chama o espírito do sistema.[138]

Há, como visto, um amplo leque de possibilidades e de formas de reconhecimento de um direito subjetivo no âmbito de um ordenamento, que vão muito além da constatação da existência de um direito a partir de mera previsão legal expressa. Em grande medida, ao se discutir as formas pelas quais seria possível reconhecer um direito, estamos a tratar, em última análise, da questão do fundamento dos direitos. Essa questão é profundamente explorada por Bobbio em "A Era dos Direitos", de 1990. Nesta obra, ao discorrer sobre "Os fundamentos dos direitos do homem", Bobbio busca responder a três questões: (i) qual a essência do problema acerca do fundamento absoluto dos direitos do homem; (ii) se um fundamento absoluto (ou "irresistível") seria possível; e (iii) se, sendo possível, seria também desejável.[139] Quanto à primeira questão, assevera:

> O problema do fundamento de um direito apresenta-se diferentemente conforme se trate de buscar o fundamento de um direito que se tem ou de um direito que se gostaria de ter. No primeiro caso, investigo no ordenamento jurídico positivo, do qual faço parte como titular de direitos e de deveres, se há uma norma válida que o reconheça e qual é essa norma; *no segundo caso, tentarei buscar boas razões para defender a legitimidade do direito em questão e para convencer o maior número possível de pessoas (sobretudo as que detêm o poder direto ou indireto de produzir normas válidas naquele ordenamento) a reconhecê-lo.*[140] (grifo nosso)

Nessa tentativa de busca por um fundamento apto a sustentar a existência de um direito que se gostaria de ter, argumenta o autor, "[...] nasce a ilusão do fundamento absoluto, ou seja, a ilusão de que de tanto acumular e elaborar razões e argumentos — terminaremos por encontrar a razão e o argumento irresistível, ao qual

---

[138] BOBBIO, N. *Teoria do ordenamento jurídico*. Apresentação: Tércio Sampaio Ferraz Júnior; Tradução: Maria Celeste C. J. Santos; Rev. téc.: Cláudio de Cicco. 6. ed. Brasília: Editora Universidade de Brasília, 1995. p. 159.
[139] Ibidem, p. 12.
[140] BOBBIO, N. Sobre os fundamentos dos direitos do homem. *In:* BOBBIO, N. *A Era dos Direitos*. 7. ed. Apresentação: Celso Lafer; Tradução: Carlos Nelson Coutinho. Rio de Janeiro: Elsevier, 2004.

ninguém poderá recusar a própria adesão."[141] Essa busca por um fundamento absoluto ou irresistível dos direitos foi compartilhada durante muito tempo pelos jusnaturalistas, mas, hoje, argumenta, essa busca tornou-se infundada, por quatro razões levantadas ao longo do texto: (i) é impossível atribuir um fundamento absoluto a noções de contornos imprecisos e conteúdo altamente abstrato, que acabam, em grande medida, sendo justificadas em função das visões políticas e ideológicas de cada intérprete; (ii) os direitos são historicamente relativos, possuindo uma natureza bastante contingente, pelo que se torna difícil atribuir-lhes um fundamento comum e absoluto; (iii) não é possível atribuir um fundamento absoluto a um conjunto de direitos heterogêneo, por vezes até menos colidentes e incompatíveis; e (iv) não é possível atribuir um fundamento absoluto a direitos que, por serem colidentes, precisarão em alguma medida serem relativizados e sopesados.[142]

Após expor as razões pelas quais entende haver uma impossibilidade de se atribuir um fundamento absoluto aos direitos, Bobbio passa a analisar se "[...] a busca do fundamento absoluto, ainda que coroada de sucesso, é capaz de obter o resultado esperado, ou seja, o de conseguir de modo mais rápido e eficaz o reconhecimento e a realização dos direitos do homem."[143] Nesse sentido, o autor argumenta que a existência de um consenso em torno do fundamento absoluto dos direitos não torna mais provável que estes sejam respeitados ou efetivados. Para o Bobbio, a busca por um fundamento absoluto é inócua e desprovida de sentido, pelo que propõe uma reorientação do problema em direção à questão da garantia da exequibilidade dos direitos: "Por isso, agora, não se trata tanto de buscar outras razões, ou mesmo (como querem os Jusnaturalistas redivivos) a razão das razões, mas de pôr as condições para uma mais ampla e escrupulosa realização dos direitos proclamados."[144] Neste sentido, o autor conclui o texto afirmando:

> *O problema fundamental em relação aos direitos do homem, hoje, não é tanto o de justificá-los, mas o de protegê-los*. *Trata-se de um problema não filosófico, mas político*. É inegável que existe uma crise dos fundamentos. Deve-se reconhecê-la, mas não tentar superá-la buscando outro fundamento absoluto para servir como substituto para

[141] BOBBIO, N. Sobre os fundamentos dos direitos do homem. *In:* BOBBIO, N. *A Era dos Direitos*. 7. ed. Apresentação: Celso Lafer; Tradução: Carlos Nelson Coutinho. Rio de Janeiro: Elsevier, 2004.
[142] BOBBIO, N. Sobre os fundamentos dos direitos do homem. *In:* BOBBIO, N. *A Era dos Direitos*. 7. ed. Apresentação: Celso Lafer; Tradução: Carlos Nelson Coutinho. Rio de Janeiro: Elsevier, 2004. p. 12-15.
[143] Ibidem, p. 15.
[144] BOBBIO, N. Sobre os fundamentos dos direitos do homem. *In:* BOBBIO, N. *A Era dos Direitos*. 7. ed. Apresentação: Celso Lafer; Tradução: Carlos Nelson Coutinho. Rio de Janeiro: Elsevier, 2004.

> o que se perdeu. Nossa tarefa, hoje, é muito mais modesta, embora também mais difícil. *Não se trata de encontrar o fundamento absoluto — empreendimento sublime, porém desesperado —, mas de buscar, em cada caso concreto, os vários fundamentos possíveis*. Mas também essa busca dos fundamentos possíveis — empreendimento legítimo e não destinado, como o outro, ao fracasso — não terá nenhuma importância histórica se não for acompanhada pelo estudo das condições, dos meios e das situações nas quais este ou aquele direito pode ser realizado.[145] (grifo nosso)

Como bem pontuado pelo autor, o esforço em torno do reconhecimento de um direito subjetivo não depende apenas da identificação de um fundamento que o sustente, ou dos vários fundamentos possíveis e aptos a sustentar a sua existência, mas do exame das circunstâncias históricas, técnicas e políticas que não apenas tornam possível a sua existência, justificam elas próprias a necessidade do reconhecimento de um novo direito. Nesse sentido, assinala Bobbio:

> *Mais uma prova, se isso ainda fosse necessário, de que os direitos não nascem todos de uma vez. Nascem quando devem ou podem nascer. Nascem quando o aumento do poder do homem sobre o homem — que acompanha inevitavelmente o progresso técnico, isto é, o progresso da capacidade do homem de dominar a natureza e os outros homens — ou cria novas ameaças à liberdade do indivíduo ou permite novos remédios para as suas indigências*: ameaças que são enfrentadas através de demandas de limitações do poder; remédios que são providenciados através da exigência de que o mesmo poder intervenha de modo protetor. [146](grifo nosso)

Essa visão histórica e dinâmica acerca da necessidade do reconhecimento de novos direitos encontra bastante abertura na cláusula geral dos direitos da personalidade no direito brasileiro, conforme veremos a seguir. Nesse sentido, Bruno Bioni argumenta que o rol de direitos da personalidade previsto nos arts. 11 a 21 do Código Civil brasileiro não é exaustivo, sendo possível reconhecer a proteção de dados pessoais como uma nova espécie de direito da personalidade.[147] Isso porque a função dos direitos da personalidade é proteger a pessoa humana, e, sendo assim, é necessário encarar a tutela dos direitos da personalidade de forma dinâmica, sendo sempre necessário revisitar e aperfeiçoar essa tutela em face de novos contextos,

---

[145] BOBBIO, N. Sobre os fundamentos dos direitos do homem. *In:* BOBBIO, N. *A Era dos Direitos*. 7. ed. Apresentação: Celso Lafer; Tradução: Carlos Nelson Coutinho. Rio de Janeiro: Elsevier, 2004. p. 16.

[146] Ibidem, p. 9.

[147] BIONI, B. R. *Proteção de dados pessoais*: a função e os limites do consentimento. 2. ed. Rio de Janeiro: Forense, 2021. p. 52.

como no caso dos desafios trazidos pelas novas tecnologias.[148]

Conforme pontuamos inicialmente, há ao menos três caminhos que podem ser trilhados no esforço de reconhecimento de um direito: i) direitos subjetivos “nascem” a partir de explícita previsão legal; ii) direitos subjetivos/princípios podem ser extraídos de um conjunto de normativas que disciplinam direitos mais amplos, a partir de uma interpretação teleológica e sistemática; e iii) direitos subjetivos podem ser derivados de valores morais, de acordo com determinadas concepções da ordem da Ética, Filosofia Política e Filosofia do Direito e de acordo com os fundamentos de legitimidade da própria ordem jurídica.

Dado que a primeira abordagem não é possível, diante da ausência de previsão legal expressa do direito à explicação na legislação pátria, resta percorrer as outras a fim de se construir um molde sólido para fundamentar a tese. A análise acerca das possíveis bases de sustentação do direito à explicação em diferentes normativas internacionais e nacionais, com destaque para a própria Lei Geral de Proteção de Dados, em comparação com regulamentos e normas estrangeiros, será objeto de discussão apresentada no capítulo seguinte. Nas subseções tratadas a seguir, apresentaremos uma fundamentação jusfilosófica baseada na dignidade e autonomia como princípios dos quais o direito à explicação é um corolário e porque deve, nesse sentido, ser reconhecido.

2.2.1 O direito à explicação como direito moral de qualquer ser humano

O reconhecimento do direito à explicação na LGPD é um preenchimento de uma lacuna no direito brasileiro, sinalizada por dispositivos e decisões esparsas, mas que até então prescindiam de um instituto que desse coerência geral. Não há na lei a previsão de um direito amplo à explicação formal regendo todas as tomadas de decisão automatizadas. Contudo, percebemos uma série de instâncias em que esse direito foi posto ou afirmado pelas cortes. Alinhando essas disposições, os princípios do ordenamento brasileiro e algumas leis setoriais, vemos uma convergência temática em volta da transparência e explicação.

O dinamismo do ordenamento jurídico é premissa para o surgimento de qualquer direito, assim como é justa a expectativa de que o sistema ofereça

[148] Ibidem, p. 55.

qualificação jurídica para todo e qualquer comportamento[149]. Assim, pode haver um conflito entre a resposta oferecida pela lei a um caso concreto e a conduta desejável. No caso em apreço, não há por que se pensar que uma pessoa deveria se ver amparada de um direito à explicação somente em decisões sobre *credit scoring*, na esfera consumerista, e não em outros contextos. A expectativa em torno de um direito à explicação amplo e abrangente é fruto de uma disposição geral do ordenamento jurídico, tendo em vista elementos basilares que sustentam o ordenamento como um todo. Aqui, principalmente, trata-se da dignidade da pessoa humana.

O primeiro passo de reconhecimento de um direito à explicação é resultado de uma evolução na compreensão social e jurídica do papel da tecnologia na sociedade. Esse ponto foi argumentado exaustivamente no primeiro capítulo dessa obra. Neste contexto, reconhece-se que a tecnologia não toma a decisão de forma completamente autônoma, e que há interesses humanos que se articulam por meio da tecnologia para executar uma decisão ou uma tomada de decisões. Esse ângulo nos permite delinear melhor a diferença entre o que é fato jurídico, como a existência de um dado pessoal, do que é ato jurídico, que são as decisões tomadas usando essas informações[150]. De um lado, a existência de dados pessoais gera direitos e deveres, enquanto atos jurídicos podem desencadear cadeias de responsabilidade.

Nesta seção, discute-se o direito à explicação como uma decorrência do processo de desenvolvimento do conceito de autodeterminação informacional, que tem por base a dignidade humana e a proteção da personalidade. Dessa premissa, extrai-se que a explicação surge como um instrumento de isonomia, e não como uma mera garantia jus processual.

2.2.1.1 A identidade e a personalidade no mundo digital

O conceito de autodeterminação surge do reconhecimento de que as interações do mundo digital forçam um reducionismo inevitável na capacidade de representar a realidade. Esse reducionismo ocorre no momento em que vidas humanas passam a ser representadas através de dados, pequenos fragmentos de informação que são tragáveis por um computador.

---

[149] KELSEN, H. *Pure theory of law*. Union, N.J: Lawbook Exchange, 1967.; e FERRAZ JÚNIOR, T. S. *Introdução ao estudo do direito, técnica, decisão, dominação*. São Paulo: Atlas, 2013.

[150] BIONI, B. R. *Proteção de dados pessoais:* a função e os limites do consentimento. 1. ed. Rio de Janeiro: Forense, 2019.

Essa explicação foi ricamente desenvolvida por Friedler *et al*.[151], que relata o processo de tomada de decisões do computador como uma série de representações e distorções que trazem um fenômeno mensurável de uma realidade infinitamente complexa para um conjunto de dados e depois regras operáveis por um computador.

Luciano Floridi[152] traz esse debate para a esfera da identidade, se interrogando sobre o que constitui a identidade individual e como ela se traduz no âmbito digital. Os dados pessoais não são, portanto, apenas uma projeção da personalidade no campo digital, mas uma projeção de apenas uma faceta de infinitas possíveis. A universalidade da identidade não consegue ser capturada e traduzida por essas migalhas representativas, por mais inúmeras que elas sejam ou estatisticamente relevantes que sejam as suas inferências.

Assim, o indivíduo não consegue ser plenamente representado com a sua identidade completa[153]. Apenas os dados servem de substrato para o processamento da máquina e não há espaço, ou meio, de apresentar o ser humano à máquina para a orientar de forma mais adequada visando obter decisões mais informadas. Como exemplo, temos os serviços que buscam aferir gostos e preferências do usuário, usando os mais variados dados para produzir inferências sobre a psique do usuário.

Essa limitação pode ter efeitos muito concretos na vida de uma pessoa.

As informações sobre indivíduos são, na realidade, fatos jurídicos dos quais um sem-número de inferências podem ser feitas, para produzir desde análises de créditos, gostos pessoais e inferências biométricas. O grande produto desses processamentos é, portanto, relacional. É da essência do tratamento automatizado produzir elementos não só na esfera subjetiva (individual), mas sim posicionar o indivíduo com relação a um grupo maior (intersubjetivo)[154]. Assim, a grande lógica do sistema é a criação de perfis (essencialmente estereótipos) e alocar direitos a esses perfis. A cada sistema, há um mecanismo discriminatório para determinar quem recebe quais direitos, a partir da análise de dados pessoais.

Os dados pessoais substituem o indivíduo e o representam em todas as

---

[151] FRIEDLER, S. A.; SCHEIDEGGER, C.; VENKATASUBRAMANIAN, S. On the (im)possibility of fairness. *ArXiv [cs, stat]*, 2016. Disponível em: http://arxiv.org/abs/1609.07236. Acesso em: 18 jun. 2020.

[152] FLORIDI, L. The Informational Nature of Personal Identity*. Minds and Machines*, v. 21, n. 4, p. 549–566, 2011.

[153] CHENEY-LIPPOLD, J. *We are data*: algorithms and the making of our digital selves. New York: New York University Press, 2017.

[154] BIONI, B. R. *Proteção de dados pessoais:* a função e os limites do consentimento. 1. ed. Rio de Janeiro: Forense, 2019.

decisões automatizadas que permeiam nossa vida em uma sociedade digital. Sobretudo, se considerarmos que sistemas algorítmicos são, por excelência, mecanismos de discriminação de dados, cabendo aos operadores o dever jurídico de não permitirem discriminações ilegais. O grande temor é que as trocas entre os espaços reais e digitais distanciam os sujeitos de direito dos seus direitos, ou que vejam a sua capacidade de influir nas decisões que lhe concernem diminuída.

É certo que esse temor nunca foi inteiramente vencido, mas encontramos no ordenamento certas garantias que dão à pessoa mecanismos de defesa contra essa alienação que pode ocorrer no mundo jurídico. A própria figura da pessoa como um instituto no direito brasileiro nada mais é que uma representação de uma personalidade real no sistema jurídico. Esse distanciamento do sujeito de direito (pessoa natural) e a sua representação digital parece ainda maior.

A preocupação com os efeitos dos tratamentos de dados é tão intrínseca à proteção de dados, que a própria LGPD tem em suas definições a figura do dado pessoal sensível. Um dado pessoal que, a partir de elementos contextuais próprios, exige um grau de proteção jurídica elevado, por poder desencadear discriminações ilegais altamente lesivas segundo a pirâmide de valores jurídicos. Por isso, a lei intercede com um mecanismo de garantias ao indivíduo e de aproximação do sujeito de direito, das informações e, portanto, das decisões que o concernem.

Portanto, proteger os dados pessoais em ambientes digitais exorbitou a percepção de privacidade como um simples direito negativo, de se proteger contra as ingerências do Estado e do mundo, como conceberam Warren & Brandeis[155]. A privacidade que versa sobre a representação de um indivíduo no ambiente digital, tem por consequência desfechos e impactos reais para a vida da pessoa. Portanto, os interesses em jogo passam a ser todos aqueles que são intermediados pelo digital. Ter capacidade de controlar suas informações virou por excelência uma proteção da autonomia individual. Nesse sentido, surge a necessidade de um direito positivo, que ofereça os meios do sujeito de direito de controlar esses fluxos informacionais que impactam a sua própria vida[156].

Mais que um dispositivo para proteção de dados pessoais, a LGPD surge com

[155] WARREN, S. D.; BRANDEIS, L. D. The Right to Privacy. *Harvard Law Review*, v. 4, n. 5, 1890. p. 193.
[156] BIONI, B. R. *Proteção de dados pessoais:* a função e os limites do consentimento. 1. ed. Rio de Janeiro: Forense, 2019.

a vocação de estabelecer os direitos da pessoa na intermediação, principalmente, com o mundo digital.

2.2.1.2 Direito à autodeterminação informacional como uma necessidade para se garantir autonomia dentro de um paradigma técnico novo

No contexto de um paradigma sociotécnico novo, caracterizado pelo crescente uso de decisões automatizadas, frequentemente cercadas de opacidade, o direito à explicação pode ser entendido como uma ferramenta capaz de atualizar o direito à autodeterminação informacional. Conforme veremos mais à frente, a explicação nada mais é que uma garantia de controle das decisões automatizadas que operam sobre dados pessoais, sendo, portanto, uma garantia à autodeterminação pessoal, atuando como instrumento essencial na articulação dos direitos da personalidade, da proteção de dados pessoais e da garantia da autonomia.

Conforme explica Rodotà, as pessoas não podem ser reduzidas a um conjunto de dados físicos e/ou virtuais, sendo sempre algo além disso[157]. Para assegurar a proteção de valores jurídicos dentro de um paradigma técnico novo, faz-se necessário o desenvolvimento de campos de interseção, que podem ser úteis num contexto permeado quase que completamente por decisões automatizadas, uma vez que talvez sejam necessários novos instrumentos para garantir direitos e princípios antigos, como o da autonomia e autodeterminação informacional.

A pessoa virtual resulta de uma interação entre sujeitos distintos do titular, na medida em que esta dispersão do ser gera a sua fragmentação que pode sofrer alterações dentro dos sistemas que interagem com estes dados. Assim, desatualizações, representações inexatas, parciais ou falsas, são comuns e colocam em risco esse ambiente em que o controle da determinação pessoal fica mais do que nunca sujeito à arquitetura e manipulação daqueles que detém o ambiente de convívio social.

Para sustentar a autodeterminação, Rodotà remonta à interação entre Biologia e Direito, em um processo recente que exigiu a redefinição do que se entende como vida e a sua valoração[158]. Uma discussão semelhante se instaura na delimitação da autodeterminação.

---

[157] RODOTÀ, S. *El derecho a tener derechos*. Madri: Editorial Trotta, S.A., 2014.
[158] RODOTÀ, S. *El derecho a tener derechos*. Madri: Editorial Trotta, S.A., 2014.

Esta alteração para o paradigma biológico permite entender a vida como uma construção da modernidade, em que é possível a autodeterminação. Rodotà, no entanto, aponta como essa concepção nova carrega em si um potencial reducionismo que a visão biológica pode implicar sobre a vida. O biodireito é um meio de colocar o direito como um obstáculo a este reducionismo científico que o preocupa e valoriza a pessoa. O biológico não pode se sobrepor ou reduzir o biográfico, e ressalta que, apesar de a existência se desenvolver em um novo contexto criado pelas descobertas biológicas, ela não pode ser resumida a um dado biológico, é a vida “pós-genômica”. E assim há a reconstrução do paradigma biológico, o determinismo não é mais a pauta, sendo substituído por um espaço decisório para a pessoa, tendo-se em vista que a sua vontade influi sobre a sua vida.

A preservação do biológico, biográfico e o digital é o processo que a autodeterminação informacional inaugura, ramo de conciliação no cruzamento de valores sociais que precisam ser equilibrados.

Seguindo o raciocínio de Rodotà, é necessário observar as alterações na distribuição de poder e a construção do ambiente jurídico que permitem a proteção da pessoa para que possa se desenvolver. E proteção da pessoa contra o poder governamental e contra a si própria, porque a sua individualização depende da responsabilidade conferida aos sujeitos, públicos e privados, que têm o dever de respeitar a sua autodeterminação. A individualização não é, portanto, um processo de separação ou isolamento dos indivíduos, como se pode pensar.

Nesse viés, para o reconhecimento da autonomia da pessoa, deve-se mapear como historicamente o direito referente à sua vida é juridicamente avaliado e desenvolvido. A sua origem normativa está na Carta das Nações Unidas, de 1945, atrelada à noção de povo, à "autodeterminação dos povos”, é subsequentemente trazida pelo Pacto Internacional sobre Direitos Econômicos, Sociais e Culturais, de 1966, bem como na Declaração Universal dos Direitos dos Povos, de 1976, e representa a necessidade de reorganizar as relações políticas em um mundo antes pautado no colonialismo[159]. Contudo, essa garantia de autonomia aos povos representa também a necessidade de autonomia do indivíduo.

Esta autonomia individual, configurada na autodeterminação da pessoa, aparece também na decisão nº 438, de 2008, do Tribunal Constitucional Italiano, como

[159] RODOTÀ, S. *El derecho a tener derechos*. Madri: Editorial Trotta, S.A., 2014.

decorrência de outros direitos constitucionais, em discussão atrelada a questões de saúde[160]. Dentre eles está a liberdade pessoal. Este direito origina-se na Carta Magna, de 1215, no que tange ao homem livre, ao *habeas corpus*. Representa a negociação feita que resulta na autolimitação estatal soberana que impede a violação da esfera pessoal, a vontade soberana deve obedecer ao que impõe a lei.

Observa-se aqui a relação íntima entre a proteção da esfera individual da ingerência externa, estatal no caso, como forma de assegurar a autodeterminação. Recorda-se, portanto, que mesmo em sua origem histórica a capacidade de delimitar espaços de interferência está diretamente ligada à capacidade de controlar os rumos da própria vida, assim como controlar os impactos da ação externa na própria vida. A privacidade deve ser entendida como uma ferramenta para o indivíduo navegar o mundo externo, na sua miríade de arranjos sociais, controlando a exposição das informações ao seu respeito e com isso modulando os impactos que o mundo externo consegue ter sobre sua autonomia[161].

Neste sentido, Rodotà apresenta a ideia de que a personalidade humana representa um limite ainda mais forte que a dignidade, tendo em vista que a dignidade tornava necessário trazer à tona o respeito à pessoa como um todo e não apenas algum de seus atributos. Por consequência, a definir sobre sua saúde, sobre sua vida, quem é o soberano é a pessoa.

A ideia de senhores de seus corpos altera-se, posteriormente, para a de senhores de suas informações, representando uma nova alteração da distribuição de poder e uma nova subjetividade. A Constituição Alemã, em 1983, passou a prever em seu texto a autodeterminação informativa como direito fundamental. Assim o poder é atribuído diretamente à pessoa e esta passa a ser um sujeito social.

Referida alteração que envolve a valorização da vontade pessoal sobre a vida consiste em uma mudança no paradigma jurídico, uma vez que no direito civil a vontade, a autonomia, está associada ao patrimônio e a realização de negócios jurídicos, há, portanto, uma dimensão econômica neste modelo. Entretanto, o que vemos com o reconhecimento da autodeterminação é que o sujeito não é apenas um agente econômico, mas é o construtor de sua própria personalidade, e por esta razão irá buscar proteger o desenvolvimento de sua vida como um todo e não apenas sua

---

[160] RODOTÀ, S. *El derecho a tener derechos*. Madri: Editorial Trotta, S.A., 2014.
[161] BIONI, B. R. *Proteção de dados pessoais:* a função e os limites do consentimento. 1. ed. Rio de Janeiro: Forense, 2019.

parcela patrimonial. Ainda que parte das ideias de autodeterminação tenham se originado na visão civilista, patrimonial, em que o indivíduo era capaz de buscar seus bens e serviços, a contextualização do termo neste meio parece equivocada, e é alterada quando é colocado no patamar de direito fundamental. Isto se nota em decisões do tribunal constitucional italiano e do tribunal europeu que nunca irão invocar normas referentes ao patrimônio para abordar o direito à autodeterminação.

A autodeterminação, portanto, trata-se de um poder individual, que limita qualquer outro poder que possa influir sobre a pessoa. E, nessa linha, que se caracteriza o valor fundacional do consentimento informado, na medida que é por meio dele que se pode expressar direitos fundamentais.

2.2.1.3 Direito à explicação como decorrência do direito à autodeterminação informativa: colocando em perspectiva o regime de direito privado e a jurisprudência constitucional brasileira

No ordenamento brasileiro, há uma relação entre a Constituição e o Código Civil, que é fonte para a existência dos direitos da personalidade, trazendo um conjunto de direitos fundamentais e um arcabouço de proteções que acompanham esses dispositivos. Essa relação entre as normas estabelece um conjunto de proteções que extrapolam a esfera patrimonial, e estendem-se a todo o desenvolvimento do indivíduo, à sua honra e ao seu bem-estar psicofísico.

Assim, é imprescindível que o direito à explicação não seja compreendido como um direito meramente patrimonial, como se verá a seguir. Diferentemente das proteções estabelecidas anteriormente, como na esfera consumerista, a explicação acompanha a proteção de dados como um direito ligado à pessoa e à sua identidade, com toda complexidade que a acompanha. Essa concepção do direito à proteção de dados como consequência e meio para o exercício da autonomia informacional do indivíduo e como direito intimamente relacionado à proteção da pessoa e de sua identidade aparece na experiência brasileira ainda em 1995, com o julgamento do RE nº 22.337-8/RS[162], primeira decisão judicial no Brasil a se valer do termo autodeterminação informacional, à época transplantada da doutrina alemã é invocada para fundamentar o cancelamento do registro negativo em banco de dados após o decurso do prazo legal. Não obstante, tal compreensão vem sendo crescentemente

[162] BRASIL. Superior Tribunal de Justiça. *Recurso Especial nº 22.337 – RS*. Relator: Ruy Rosado Aguiar. Data de Publicação: 20/03/1995. *In: R. Sup. Trib. Just. Brasília*, a.8 (77), jan. 1996.

reforçada pelo judiciário brasileiro, como se observa da decisão conferida em sede de medida cautelar no âmbito da ADI 6.387/DF, na qual a Ministra Rosa Weber reconhece a existência de um direito fundamental autônomo à proteção de dados pessoais, com clara ênfase na garantia da autodeterminação informativa do titular[163]. No racional desenvolvido na decisão, o direito fundamental à proteção de dados pessoais pode ser derivado de uma leitura sistemática do texto constitucional, sobretudo a partir do direito fundamental à dignidade da pessoa humana (art. 1º, III), de uma leitura contemporaneamente e contextualmente situada da proteção constitucional à intimidade (art. 5º, X) e do reconhecimento do *habeas data* como instrumento central na tutela do direito à autodeterminação informativa (art. 5º, LXXII). Conforme assinalam Mendes, Rodrigues Júnior e Fonseca, a decisão do STF na ADI 6.387/DF estabelece um verdadeiro paradigma na jurisprudência constitucional brasileira em matéria de proteção de dados pessoais:

> O significado histórico da decisão do STF pode ser equiparado ao clássico julgamento do Tribunal Constitucional Federal alemão, em 1983, relativamente à Lei do Recenseamento. Ao fazer referência ao julgado, o STF expressamente mencionou o conceito de *autodeterminação informativa*, já positivado na Lei n. 13.709/2018 (Lei Geral de Proteção de Dados), a fim de ressaltar o necessário protagonismo exercido pelo cidadão no controle do que é feito com seus dados. Assim, pôs-se em destaque a existência de finalidades legítimas para seu processamento, bem como da necessidade de implementação de medidas de segurança para tanto. Segundo o Ministro Gilmar Mendes, o conteúdo desse direito fundamental exorbita àquele protegido pelo direito à privacidade, pois não se limita apenas aos dados íntimos ou privados. Ao contrário, refere-se a qualquer dado que possa identificar um indivíduo.[164]

É nesse sentido que a visão da explicação somente como um regime de responsabilização por dano não é uma figura suficiente para contemplar todos os interesses da pessoa nos tratamentos automatizados. Essa constatação surge inclusive do regime escolhido pela LGPD para estabelecer direitos do titular de dados.

O regime patrimonial permite a disposição plena do proprietário, em um direito

---

[163] BRASIL. Supremo Tribunal Federal. *Medida Cautelar na Ação Direta de Inconstitucionalidade nº 6.387 Distrito Federal*. Relator: Min. Rosa Weber. Data de Julgamento: 24/04/2020. Disponível em: http://www.stf.jus.br/arquivo/cms/noticiaNoticiaStf/anexo/ADI6387MC.pdf. Acesso em: 25 nov. 2020.

[164] MENDES, L. S.; RODRIGUES JÚNIOR, O. L.; FONSECA, G. C. S. da. O Supremo Tribunal Federal e a Proteção Constitucional dos Dados Pessoais: Rumo a um Direito Fundamental Autônomo. *In*: MENDES, L. S.; DONEDA, D.; SARLET, I. W.; RODRIGUES JÚNIOR, O. L.; BIONI, B. (Coords.). *Tratado de Proteção de Dados Pessoais*. Rio de Janeiro: Forense, 2021. p. 67.

que é oponível contra a sociedade toda e sujeito ao arbítrio do seu detentor[165]. Seguindo raciocínio aqui estabelecido, é impossível conciliar a abordagem apenas patrimonial com a plena proteção da dignidade humana, visto que o regime patrimonial sujeitaria o indivíduo a relações não isonômicas de troca, entregando direitos sobre a sua personalidade (os dados) na aquisição de bens e serviços[166]. É por isso que o legislador preferiu não se apoiar apenas nessa esfera, que em uma lógica de sistemas automatizados apenas perpetuaria a assimetria entre o titular e os detentores do sistema.

Se observarmos as disposições presentes no Código Civil brasileiro, os direitos da personalidade são sistematizados com o objetivo de distingui-los dos direitos subjetivos, recebendo características como sua intransmissibilidade e irrenunciabilidade (art. 11). Como explica Doneda[167], não é funcional igualar os dois tipos de direitos, além de que a cláusula geral já possibilita esta distinção. Portanto, a tutela da personalidade não pode ser limitada por atos ordinários, uma vez que é estabelecida constitucionalmente. Por sua generalidade, a proteção conferida por ela é integral, aplicável a qualquer situação. Garantindo, e mesmo ampliando, essa tutela, o art. 12, do CC, prevê a responsabilidade civil e a legislação também estabelece outras sanções.

O texto normativo também apresenta proteção à integridade psicofísica, protegidas de forma conjunta (art. 13 a 15). O autor aponta que sobre este tema ainda existem controvérsias quanto à aplicação do direito à personalidade, utilizando o exemplo da cirurgia transexual, que tem como função primárias o favorecimento do desenvolvimento da personalidade. No entanto, há controvérsias sobre a interpretação do dispositivo ou também sobre a questão das diretivas antecipadas de vontade. Dessa forma, destaca-se também o papel da interpretação ampla da norma, para evitar uma restrição sobre a personalidade. O código também dá grande importância para o direito ao nome, regulado de forma a evidenciar sua relação com a formulação de um direito à identidade pessoal, propriamente dito.

Por fim, há a proteção da imagem e da honra nos artigos subsequentes, que

---

[165] TEPEDINO, G. *et al.* (Ed.). *Lei Geral de Proteção de Dados Pessoais e suas repercussões no direito brasileiro*. São Paulo: Thomson Reuters, Revista dos Tribunais, 2019.
[166] DONEDA, D. *Da privacidade à proteção de dados pessoais*. 7. ed. São Paulo: Revista dos Tribunais, 2019.
[167] DONEDA, D. *Da privacidade à proteção de dados pessoais*. 7. ed. São Paulo: Revista dos Tribunais, 2019.

também são direitos extrapatrimoniais da pessoa. Há ponderações entre o direito de preservação da pessoa em si e o direito à informação, limitando-se apenas em certas ocasiões as divulgações de nome e imagem, muito pautada na permissão quando não há violação à honra. Por mais que essas proteções não se confundam com o regime de privacidade, atrelando-se, eles mostram como a proteção da personalidade é um conjunto amplo de institutos que vai abarcar as múltiplas facetas que integram a liberdade de autodeterminação.

Daí a segunda razão para o abandono da abordagem patrimonial, e também outra razão para se integrar a explicação como um direito integrante da personalidade, e não mero mecanismo de correção incidental sobre relações negociais. O viés negocial privilegia relações e desigualdades patrimoniais existentes na sociedade, o que compromete a igualdade entre as pessoas e inclusive fere princípios básicos do sistema democrático[168].

A aproximação da proteção de dados pessoais a um mecanismo de igualdade, e não apenas de privacidade e discrição de um usuário de internet, eleva a importância do instituto e exige proteções jurídicas superiores àquelas da propriedade. Afasta-se por completo os dados pessoais da figura de um bem patrimonial.

Isso não significa que dados pessoais não sejam mercantilizáveis. Não são raras, inclusive, as associações, frequentemente equivocadas, entre dados pessoais e matérias primas (*data is the new oil*).[169] É evidente que há um interesse econômico enorme na economia dos dados, a grande força motriz da inovação digital que vivemos hoje. Na verdade, o que se defende é que a titularidade dos dados pessoais não pode ser reduzida apenas à sua dimensão patrimonial, devendo todos que tratarem dados pessoais respeitar sua importante dimensão extrapatrimonial.

Ou seja, dentro da Constituição de 1988, o regime de titularidade é um gênero amplo, dentro do qual a espécie de propriedade se encontra. Trata-se de uma relação e direito imediato, atribuída a alguém[170]. Direito esse que é imediatamente associado à personalidade, segundo o próprio texto constitucional. Portanto, a proteção aos dados pessoais trata da uma expansão da esfera da proteção da personalidade para

---

[168] TEPEDINO, G. *et al.* (Ed.). *Lei Geral de Proteção de Dados Pessoais e suas repercussões no direito brasileiro*. São Paulo: Thomson Reuters, Revista dos Tribunais, 2019.
[169] RAMGE, T.; MAYER-SCHONBERGER, V. *Reinventing Capitalism in the Age of Big Data*. New York: Basic Books, 2018.
[170] TEPEDINO, G. *et al.* (Ed.). *Lei Geral de Proteção de Dados Pessoais e suas repercussões no direito brasileiro*. São Paulo: Thomson Reuters, Revista dos Tribunais, 2019.

um eixo onde antes não havia. Ganha-se uma tutela relacional cuja elasticidade acompanha os usos de dados pessoais e os direitos dos titulares[171].

Esse campo da tutela relacional é a essência do direito à explicação. A explicação nada mais é que uma garantia de controle das decisões automatizadas que operam sobre dados pessoais, sendo, portanto, uma garantia à autodeterminação pessoal. Por isso, é elemento essencial de coerência desses diversos aspectos da personalidade, da proteção de dados e da autonomia.

Tanto o seu objeto é uma garantia de controle que o art. 20 da LGPD, explorado a fundo no subcapítulo 3.2.2, garante o direito de revisão a todos os titulares de dados, sem condicioná-lo a nenhuma base legal, ou mesmo a fazer restrição do campo de aplicação. *In verbis*:

> Art. 20. *O titular dos 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 destinadas a definir o seu perfil pessoal, profissional, de consumo e de crédito ou os aspectos de sua personalidade. (grifo nosso).

Verifica-se que as hipóteses do *caput* são meramente exemplificativas, pelo emprego da palavra “incluídas”, e não outro termo que aponte que o rol é taxativo. Verificamos que os exemplos a seguir são as práticas setoriais, como o perfilamento pessoal de consumo ou de crédito (*grouping* ou *profiling*), práticas que já encontravam, de certo modo, mesmo que não expressamente, tutela de explicação no direito brasileiro. Há, portanto, o nítido objetivo de se expandir esse mesmo direito para além das práticas já previstas, no intuito de trazer essa garantia para todo o escopo de aplicação de dados, que é amplo.

Essa análise hermenêutica se justifica pelo raciocínio que se desenvolveu nesta seção, mostrando que a explicação é uma peça de um mosaico, que envolve diversos aspectos da personalidade. O art. 20 da LGPD é o ponto de encontro que preenche uma grande lacuna do ordenamento jurídico, e o próprio texto legal aponta para a sua intenção de complementar as disposições existentes com o rol exemplificativo.

Esse entendimento vem também sendo corroborado por decisões do STJ, principalmente na esfera consumerista. Em decisão recente, o Ministro Gilmar

---

[171] BIONI, B. R. *Proteção de dados pessoais:* a função e os limites do consentimento. 1. ed. Rio de Janeiro: Forense, 2019.

Mendes defendeu a figura do “devido processo informacional” em sua decisão. Chamamos particular atenção para o seguinte excerto de seu voto, na ADI nº 6.389:

> É possível identificar como corolário da dimensão subjetiva do direito à proteção de dados pessoais, a preservação de verdadeiro “devido processo informacional” (*informational due process privacy right*), voltado a conferir ao indivíduo o direito de evitar exposições de seus dados sem possibilidades mínimas de controle, sobretudo em relação a práticas de tratamento de dados capazes de sujeitar o indivíduo a julgamentos preditivos e peremptórios.[172]

O Ministro afirma inequivocamente a evolução aqui defendida. Do instituto da privacidade e da proteção da autonomia privada surge um direito também à isonomia da informação no campo das decisões automatizadas. Dessa isonomia, depreende-se também o direito de se ter um controle humano no processo, o chamado *“*human-in-the-loop*”*, até nos casos em que o tratamento automatizado motive uma decisão humana[173].

O devido processo informacional não se restringe a uma aplicação limitada de privacidade, mas serve um propósito de simetria entre os indivíduos de uma sociedade[174]. Nesse sentido, a simetria cumpre um papel fundamental ao proteger, além de direitos individuais, um interesse coletivo no uso da tecnologia como um todo: a preservação da confiança.

A confiança é o elemento crucial para o funcionamento da economia digital, reconhecido em diversas instâncias nacionais e internacionais, como na decisão da Corte Constitucional alemã sobre a Lei do Recenseamento (*Vokszahlungesetzt*) de 1983[175], do Affaire Safari, que culminou na Lei de Informática de Liberdades francesa, em 1978[176], ou mesmo nas diretrizes da OCDE sobre o tratamento de dados pessoais. A LGPD vem na esteira desse processo histórico, fortalecendo o direito da personalidade (art. 1º) e da inovação econômico-tecnológica (art. 2º) nesse campo de

---

[172] BRASIL. Supremo Tribunal Federal. *ADI nº 6.389/DF*. Relator: Min. Rosa Weber. Data de Julgamento: 26/11/2020. Data de Publicação: 30/11/2020. Disponível em: http://portal.stf.jus.br/processos/detalhe.asp?incidente=5895168. Acesso em: 24 fev. 2021.

[173] BIONI, B.; MARTINS, P. Devido processo informacional: um salto teórico-dogmático necessário? *Jota*, 15 jul. 2020. Disponível em: https://www.jota.info/opiniao-e-analise/artigos/devido-processo-informacional-um-salto-teorico-dogmatico-necessario-15072020. Acesso em: 15 jul. 2020.

[174] BIONI, B.; MARTINS, P. Devido processo informacional: um salto teórico-dogmático necessário? *Jota*, 15 jul. 2020. Disponível em: https://www.jota.info/opiniao-e-analise/artigos/devido-processo-informacional-um-salto-teorico-dogmatico-necessario-15072020. Acesso em: 15 jul. 2020.

[175] BIONI, B. R. *Proteção de dados pessoais:* a função e os limites do consentimento. 1. ed. Rio de Janeiro: Forense, 2019.

[176] BOUCHER, P. “Safari” ou la chasse aux Français. *Le Monde*, p. 9, 21 mar. 1974. Disponível em: https://www.cnil.fr/sites/default/files/atoms/files/le_monde_0.pdf. Acesso em: 21 jun. 2021.

interseção. O art. 20 oferece, inequivocamente, a sustentação para um direito amplo de devido processo informacional e assegura a simetria entre os titulares de dados e os processadores no que aqui chamamos de Direito à Explicação.

2.2.2 Direito à explicação como garantia de decisões mais justas e adequadas

Uma forma de entender a zona de interseção entre algoritmos e concepções de justiça pode ser ilustrada da seguinte forma: se algoritmos estão mediando as nossas operações no meio digital, eles estão necessariamente modelando o gozo de todo e qualquer direito que passe por um meio digital, direitos esses que podem ser de ordem privada e contratual, mas também direitos advindos de normas de todos os níveis. Apesar do otimismo em torno dos sistemas automatizados e da tão propalada eficiência, neutralidade e objetividade que eles seriam capazes de proporcionar, faz-se necessário analisá-los criticamente, para além do discurso ingênuo e otimista. Se é verdade que algoritmos e sistemas automatizados podem promover justiça e intermediar a alocação de direitos, é também possível que, como produtos da razão e da cultura humana, eles também possam reproduzir, amplificar e atualizar desigualdades e injustiças já existentes, conforme aponta Virginia Eubanks, em seu seminal "*Automating inequality: how high-tech tools profile, police, and punish the poor*".[177] Nesse contexto, o direito à explicação pode ser entendido como uma forma de garantir que as decisões automatizadas operem de forma mais justa, transparente e adequada, permitindo aos indivíduos entenderem como esses sistemas os afetam e serem capazes de exercerem sua autodeterminação informacional. A seguir, apresentamos alguns exemplos nos quais é possível observar como algoritmos passaram a atuar como mecanismos de intermediação e alocação de direitos, bem como alguns dos problemas presentes nesse processo.

O Waze, serviço de rota da Google, orienta os usuários (e motoristas) a pegarem as rotas mais rápidas. O Waze está fazendo isso com centenas de milhares de usuários simultaneamente. Sendo assim, o serviço tem a capacidade de alocar,

[177] EUBANKS, Virginia. *Automating inequality: how high-tech tools profile, police, and punish the poor*. New York: St Martin's Press, 2018.

através da indução de comportamento (*nudging*)[178], o fluxo do trânsito na cidade.[179] Por mais absurdo que pareça, vale olhar o caso de um cidadão que decidiu usar 100 celulares para gerar uma falsa indicação de trânsito no Google Maps. Diante do tráfego congestionado, o serviço acaba redirecionando os usuários para outras rotas. O Waze opera da mesma forma.

O serviço é, sem dúvida, útil, tanto no plano individual, onde o usuário teoricamente acha o caminho mais curto para o seu destino, quanto no coletivo, pois consegue acomodar melhor os usuários conforme o espaço disponível nas vias da cidade. Contudo, a primeira inquietação surge ao perceber que uma empresa privada está fazendo o papel de gestão e engenharia de trânsito de uma cidade. Se parece absurdo, de novo, vale ver que a Prefeitura de São Paulo há tempos troca informações com o aplicativo Waze para poder aprimorar seus dados de trânsito.[180]

A segunda preocupação é que a empresa estaria, de fato, interferindo na liberdade de ir e vir dos cidadãos. O algoritmo poderia ser programado, por exemplo, para conduzir os usuários para uma rota, enquanto deixa outra livre para o tráfego de ambulâncias. Da mesma forma, poderia orientar usuários a seguirem por uma determinada rota para passar em frente de um estabelecimento comercial que pagou por essa forma de propaganda. Essencialmente, não se sabe qual é a conta ou os critérios que a empresa faz para sugerir que um determinado usuário vá por um caminho ou por outro. Essa ingerência no gozo de liberdades tem de ser informada e justificada perante o usuário, grupo de usuários ou a sociedade.

Não se trata de uma ingerência direta, apoiada pelo poder público, mas um poder de condução social que emana do fato de contingentes enormes de pessoas usarem o mesmo serviço, que, por sua vez, compartilham uma quantidade imensurável de seus dados pessoais. Há, portanto, um acoplamento do poder de influência desses serviços a estruturas públicas e privadas, num contexto em que os serviços algorítmicos se tornam *guardiões* ou *moderadores* (*gatekeepers*) de um

---

[178] THALER, R. H.; SUNSTEIN, C. R. *Nudge: improving decisions about health, wealth, and happiness*. New York: Penguin Books, 2009.

[179] MACHADO, C. C. V. Cidade dos algoritmos: A Ética da Informação nas Cidades Inteligentes. *Instituto de Tecnologia e Sociedade do Rio — ITS Rio*, mar. 2018. Disponível em: https://itsrio.org/wp-content/uploads/2018/03/caio_machado_etica.pdf. Acesso em: 19 abr. 2021.

[180] PREFEITURA DE SÃO PAULO. Prefeitura de São Paulo anuncia parceria com Waze. *Cidade de São Paulo*, 20 set. 2017. Disponível em: http://www.capital.sp.gov.br/noticia/prefeitura-de-sao-paulo-anuncia-parceria-com-waze/. Acesso em: 19 abr. 2021.

interesse coletivo.[181] Neste contexto, esses *gatekeepers* podem estar moderando posições centrais da alocação de direitos. No cenário brasileiro, dois exemplos significativos dessa dinâmica podem ser mencionados: a automação da gestão do acesso à saúde no estado de São Paulo, por meio da Central de Regulação de Ofertas de Serviços de Saúde (CROSS)[182] e a gestão de auxílios emergenciais no contexto da pandemia de Covid-19, com base nos dados da Dataprev.[183]

---

[181] KAK, A. Regulating Biometrics: Global Approaches and Urgent Questions. *AI Now Institute*, set. 2020. Disponível em: https://ainowinstitute.org/regulatingbiometrics.pdf. Acesso em: 20 abr. 2021. p. 31.

[182] No estado de São Paulo, a Central de Regulação de Oferta de Serviços da Saúde (CROSS) reúne informações sobre serviços hospitalares e opera a sua distribuição em toda a rede da saúde. Embora as decisões ocorram por profissionais da saúde, o sistema divide o processo decisório entre diferentes atores que atuam de forma independente. A alimentação das informações, bem como as avaliações de risco e urgência são, portanto, realizadas de forma descentralizada. A CROSS veio ao centro da discussão nos últimos meses devido ao seu papel durante a pandemia de COVID-19. É com base em seu sistema que se realizam as alocações de vagas em hospitais e leitos em UTI no estado. A CROSS permitiu que a gestão distribuísse pacientes para hospitais com disponibilidade de vagas em várias regiões do estado. A plataforma foi utilizada, ainda, para reorganizar os leitos entre hospitais exclusivos para o tratamento da COVID-19 e demais enfermidades. A centralização das decisões na CROSS na pandemia pode ainda trazer conflitos entre as gestões estaduais e municipais. Diante da escassez de vagas é possível que tal sistema tenha a capacidade de escolher quais regiões serão mais ou menos demandadas nos serviços de saúde. Em CPI realizada em 2018, questionou-se o modelo de gestão do acesso à saúde pela administração indireta. Embora a administração argumente que o sistema não seja uma fila de espera e de que as decisões são tomadas apenas por profissionais, diante da sua opacidade, é necessário problematizar a forma como o sistema informatizado pode influenciar as decisões dos médicos e criar assimetrias regionais. Em última instância, é necessário verificar como o sistema automatizado da CROSS, por meio de seus algoritmos, decide e define quais pacientes terão preferência no atendimento, onde estes serão alocados e a quais serão negados serviços básicos essenciais de saúde. Cf. SÃO PAULO. GOVERNO DO ESTADO DE SÃO PAULO. Central de Regulação de Oferta de Serviços de Saúde. Disponível em: http://www.cross.saude.sp.gov.br/. Acesso em: 20 abr. 2021; SÃO PAULO. Assembleia Legislativa do Estado de São Paulo. *CPI – Organizações Sociais da Saúde*. 07 jun. 2018. Disponível em: https://www.al.sp.gov.br/spl/2018/06/Transcricao/1000221166_1000187328_Transcricao.pdf. Acesso em: 20 abr. 2021.

[183] A pandemia de COVID-19 trouxe à tona questões importantes sobre o uso de dados pessoais pelo poder público. O tratamento dos dados pessoais também entrou em questão para operacionalizar o fornecimento do auxílio emergencial previsto na Lei nº 13.982/2020, regulamentado pelo Decreto nº 10.316/2020. Após uma intensa discussão para sua aprovação, veio à tona o problema de verificação dos dados e dos critérios necessários para o recebimento do benefício. O decreto atribui ao Ministério da Cidadania, com auxílio do Ministério da Economia, a tarefa de gerir a distribuição do benefício. A solução encontrada para o caso, nos termos da Portaria nº 351/2020, do Ministério da Cidadania, foi a contratação da Caixa Econômica Federal para a operacionalização dos pagamentos. Para tratamento dos dados necessários para a verificação dos critérios de elegibilidade a portaria elegeu a empresa pública Dataprev. No dia 7 de abril de 2020, a Caixa Econômica Federal disponibilizou o *site* e o aplicativo para celular para que os cidadãos realizassem o pedido do benefício através do fornecimento de seus dados. Esses dados, uma vez em posse da instituição financeira, são enviados à Dataprev, que realiza o processamento dos dados, nos termos do art. 6º da referida portaria. Os dados dos requerentes ao benefício são cruzados com bancos de dados de vários órgãos da administração pública. Com autorização do Decreto nº 10.316/2020, o Ministério da Cidadania transferiu à empresa as bases de dados do Cadastro Único (CadÚnico) e dos beneficiários do Bolsa Família. O dispositivo ainda autorizou a utilização das bases de dados do Ministério da Economia com informações necessárias para a verificação dos critérios de elegibilidade. Por meio do Contrato Administrativo nº 12/2020, assinado entre o Ministério da Cidadania e a empresa, foi assegurado a esta última o acesso

O que casos práticos, como o exemplo do uso do Waze, da automação da gestão de serviços de saúde no estado de São Paulo por meio da CROSS e da automação da concessão do auxílio emergencial (Dataprev) nos indicam é que algoritmos são procedimentos que podem resultar em decisões sobre a alocação de direitos. Isso significa que não só os desenvolvedores deveriam justificar a alocação de direitos escolhida, como as alocações possíveis deveriam obedecer a conceitos de justiça que são completamente incompatíveis. Por exemplo, se igualdade consiste em oferecer resultados iguais a todos os interessados, enquanto equidade consiste em oferecer resultado proporcional ao ponto de partida, algoritmos operando nesses dois

---

e interferência em quaisquer bases de dados necessárias à prestação de seus serviços. E, neste caso, o critério de necessidade carece de uma definição precisa, o que pode levar a um acesso e uso de dados em excesso ao estritamente necessário para a concessão do benefício, resultando em riscos no tratamento dos dados dos cidadãos, que, inclusive, possivelmente se encontram na camada mais vulnerável da população nacional. O primeiro problema enfrentado na verificação dos critérios de elegibilidade foi o funcionamento do aplicativo, que apresentou falhas na instrucionalidade ou no funcionamento da aplicação, ocasionando filas e aglomerações nas agências bancárias onde os valores referentes ao benefício social poderiam ser sacados. Um segundo problema ocorreu após a estabilização do aplicativo. Os problemas na análise relacionam-se principalmente à integridade dos bancos de dados públicos. Segundo o portal da Dataprev, em publicação do dia 20 de junho de 2020, são 23 bases governamentais utilizadas na avaliação dos critérios de elegibilidade, cujas informações nem sempre se encontram atualizadas. Essa questão, por si, acende um sinal de alerta pela qualidade dos sistemas informacionais das organizações públicas e indica que pode haver outras fragilidades. Nesse contexto de baixa qualidade nas informações e a quantidade de bases de dados utilizadas, cabe indagar acerca dos critérios objetivos utilizados para a concessão do benefício e como funciona o processo de revisão das decisões. Embora os órgãos garantam essa possibilidade, o procedimento não se encontra suficientemente esclarecido. O portal do Ministério da Cidadania não disponibilizou ao público o Projeto Básico, anexo do contrato administrativo firmado com a Dataprev, no qual o processo deveria estar detalhado. Além disso, cabe avaliar a explicação prestada àqueles que tiveram os benefícios negados e quais seriam as possibilidades para que possam contestar a decisão. Milhões de pedidos de benefícios foram negados sem que houvesse justificativas ou explicações suficientes. Justificativas genéricas foram entregues, sem canais para questionamentos e argumentação. Muitas vezes o sistema informava que havia incompatibilidade de informações entre os bancos de dados, sem que se explicasse quais seriam as incompatibilidades, deixando os usuários na difícil situação caso desejassem corrigir. Cf. BRASIL. *Portaria nº 351, de 7 de abril de 2020*. Ministério da Cidadania. Diário Oficial da União, Brasília-DF, Seção 1 — Extra, edição 67-B, publicado em 7 abr. 2020, p. 13. Disponível em: https://www.in.gov.br/en/web/dou/-/portaria-n-351-de-7-de-abril-de-2020-251562808. Acesso em: 24 fev. 2021; BRASIL. *Decreto nº 10.316, de 7 de abril de 2020*. Brasília: Presidência da República, 07 abr. 2020. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2019-2022/2020/decreto/d10316.htm. Acesso em: 24 fev. 2021; BRASIL. *Contrato Administrativo nº 12/2020*. Ministério da Cidadania, Secretaria Executiva, Subsecretaria de Assuntos Administrativos. Processo Administrativo nº 71000.022387/2020-55. Diário Oficial da União, Brasília-DF, Seção 3, nº 91, de 14 de maio de 2002, p. 5. Disponível em: http://www.mds.gov.br/webarquivos/acesso_informacao/contratos/2020/12.2020/Contrato%20Administrativo%20n%C2%BA%2012.2020%20-%20DATAPREV.pdf. Acesso em: 23 nov. 2020; AUXÍLIO emergencial: Bases de dados utilizadas pela Dataprev. *Dataprev*. 20 jun. 2020. Disponível: https://portal2.dataprev.gov.br/bases-de-dados-ultilizadas-no-processamento-do-auxilio-emergencial. Acesso em: 23 nov. 2020; e VELOSO, A. C; CARDOSO, L.; BRÊTAS, P. Jogo dos 7 erros: auxílio de R$ 600 é negado a quem tem requisites e concedido a quem não precisa. *O Globo*, 5 jun. 2020. Disponível em: https://oglobo.globo.com/economia/jogo-dos-7-erros-auxilio-de-600-negado-quem-tem-requisitos-concedido-quem-nao-precisa-1-24464513. Acesso em: 20 abr. 2021.

paradigmas podem oferecer resultados completamente distintos[184].

Apesar de termos a impressão de que é a tecnologia que toma essas decisões, como descreve Gillespie quando discute algoritmo como um “talismã” ou uma “sinédoque”[185] que oculta os humanos e as organizações por trás, temos que entender que algoritmos, no seu modelo tradicional, apenas operacionalizam decisões tomadas pelos seus desenvolvedores, sejam essas decisões tomadas de forma deliberada ou não deliberada, como no caso de aprendizado da máquina. Reitera-se aqui: o grande problema é que sequer o usuário tem acesso às decisões tomadas. E isto afeta não só a capacidade de perceber e responder a essas decisões, mas também influencia diretamente a legitimidade da alocação final de direitos oferecida pelo algoritmo.

O que transparece na análise sociológica de algoritmos é que eles são essencialmente procedimentos, ou seja, aliados de mecanismos de execução digital. A grande mudança jaz, principalmente, na escala e complexidade maior do que o que existia antes em âmbito mecânico. Isso significa que a sua utilização como intermediários, ou até mesmo como uma espécie de procuradores tecnológicos para a tomada de decisão, está essencialmente sujeita às mesmas avaliações éticas sobre a justiça que julgam as decisões jurídicas. Ou seja, como ressaltaram Gillespie e Cheney-Lippold[186], algoritmos nos prometem uma ideia de objetividade científica, mas em essência carregam decisões humanas.

Assim, pode-se aqui resgatar o pensamento de Rawls, que percebeu nos procedimentos o elemento essencial da concretização de valores de justiça, o que ele chamou de justiça procedimental[187]. Em essência, os conceitos de justiça podem variar, pregando valores morais dos mais diversos. Isso significa que pessoas podem se encontrar em embates éticos acerca dos valores mais adequados para determinada situação, como a distribuição equânime de recursos entre as pessoas, uma distribuição corretiva ou mesmo uma distribuição utilitária que maximize o ganho agregado coletivo.

Independente da correção da escolha moral, o balanceamento e a concretização desses valores passarão invariavelmente por um procedimento de

---

[184] FORSYTH, D. R. Conflict. *In:* FORSYTH, D. R. *Group dynamics*. 5. ed. Belmont: CA: Wadsworth, Cengage Learning, 2006.
[185] GILLESPIE, T. Algorithm*. In:* PETERS, B. *Digital Keywords*. Princeton: Princeton University Press, 2016. p. 18-30.
[186] CHENEY-LIPPOLD, J. *We are data*: algorithms and the making of our digital selves. New York: New York University Press, 2017.
[187] RAWLS, J. *Uma Teoria da Justiça*. 3. ed. São Paulo: Martins Fontes, 2008.

aplicação, que partirá dos critérios definidores do valor moral, a fim de concretizá-los, através de etapas, em uma alocação real de recursos. Voltamos à antiga anedota das crianças que disputam um pedaço de bolo e o pai que determina que um será encarregado de cortar o bolo, enquanto o segundo terá direito de escolher o pedaço.

O valor subjacente é de que ambas as crianças são iguais e que devem receber porções igualitárias do doce. Contudo, o procedimento de alocação torna-se um valor em si, a partir do momento que concede aos interessados os meios de construírem uma solução que seja aceitável para ambos e que tenha mecanismos de controle contra abusos da outra parte. Esse valor procedimental acaba preponderando em eficácia sobre o próprio valor abstrato de igualdade, pois as garantias efetivam uma distribuição parecida (portanto "justa"), ainda que potencialmente o pai tivesse a capacidade de cortar o bolo igualmente com mais precisão sozinho.

Nesse sentido, metaforicamente, é possível afirmar que um algoritmo opera o corte e a distribuição do "bolo" para os interessados do bem jurídico que ele opera.

Nos exemplos apresentados acima, sobre a alocação de recursos no campo da saúde e de auxílios do Estado, conseguimos perceber a importância da automação nos mínimos critérios. No caso do CROSS, por exemplo, observa-se que o *software* da regulação de leitos dá particular destaque a mulheres grávidas, por exemplo, para que reguladores deem especial atenção àqueles casos e os tratem mais rápido.

O princípio governante do sistema é de que todos têm direito à vida e à saúde. A escolha de dar tratamento prioritário a determinados pacientes emana de escolhas éticas mais ou menos institucionalizadas, como, por exemplo, a prioridade a mulheres grávidas. Essa escolha emana não da vida como um princípio abstrato, mas a partir da concretização desse valor, que encontra barreiras à sua execução na vida real. Barreiras essas que podem colocar os interesses individuais de dois pacientes em posição de concorrência e obrigam o poder público a tomar decisões que podem violar o direito à vida de um indivíduo ou outro. No caso do CROSS, a mera priorização de uma paciente, dada através da indicação automática do *software*, pode significar lesão aos interesses de outro paciente que se viu preterido em uma situação de urgência.

O ordenamento jurídico reconhece essas situações nas quais há um dever de maximização de valores normativos gerais, muitas vezes impossíveis de serem concretizados em conjunto diretamente a partir de sua forma abstrata. Alexy reconhece um mecanismo assim quando discute os princípios jurídicos como "comandos de maximização", ou seja, ordens que devem ser alcançadas na melhor

forma possível[188]. Esses comandos têm aplicação circunstancial e gradativa, diferentemente das regras jurídicas comuns, que têm aplicação "binária", na lógica do *tudo ou nada*, que devem ser aplicáveis ou não, sem meio termo.

Nesse sentido, o que se percebe na discussão normativa, seja ela jurídica ou não, é que a procedimentalização é o recurso que encontramos para concretizar valores normativos diversos que se encontram em concorrência no caso concreto. É, por exemplo, o grande embate moral do algoritmo subjacente ao "*The Moral Machine*", do MIT, onde o carro autônomo é obrigado a escolher entre duas situações danosas[189]. Ou, por exemplo, se o algoritmo de um sistema coletivo de roteamento de tráfego precisar reorientar os carros da cidade de São Paulo para liberar o caminho de uma ambulância.

Friedler *et al.*[190] explicam essa constatação inevitável. O uso de algoritmos opera com o objetivo de tomar dados da realidade e torná-los em decisões. Portanto, temos dois espaços: o primeiro sendo o espaço das características (*feature space*) e o segundo sendo o do resultado, que sai no campo das decisões (*decision space*). A transição entre esses dois não é nada linear, apesar de tomada como algo auto-evidente. Analisemos a seguir.

A primeira decisão vem na compreensão de quais aspectos da realidade se deseja ver representados no campo de funcionamento do algoritmo para se tomar uma decisão. Suponhamos que nós queiramos preencher uma vaga de trabalho.

Assim, é preciso conceber um construto comparativo para que se possa comparar pessoas. "Aptidão", que significa, grosso modo, a capacidade individual de se realizar alguma coisa, pode ser um construto adequado para estabelecermos um comparativo entre os candidatos. "Aptidão" não é uma característica da natureza, mas um conceito complexo e multifacetado (e mesmo sem definição clara), que reside no campo da subjetividade humana. Ainda que subjetivo, é esse o resultado essencial que justifica a escolha do candidato. Esse é o primeiro espaço, que estabelece significado das observações que seguem

Na sequência, precisamos transformar "aptidão" em uma característica

---

[188] ALEXY, R. *Teoria dos direitos fundamentais*. São Paulo: Malheiros, 2008.
[189] MIT MEDIA LAB. *The Moral Machine*. Disponível em: http://moralmachine.mit.edu. Acesso em: 29 abr. 2020.
[190] FRIEDLER, S. A.; SCHEIDEGGER, C.; VENKATASUBRAMANIAN, S. On the (im)possibility of fairness. *ArXiv [cs, stat]*, 2016. Disponível em: http://arxiv.org/abs/1609.07236. Acesso em: 18 jun. 2020.

observável. Podemos escolher de uma série de características humanas como maleabilidade, persistência, capacidade cognitiva, capacidade de trabalhar em equipe. Cada uma dessas características enseja uma metodologia de mensuração diferente, podendo ser avaliadas desde testes de quociente intelectual (QI), a execução de tarefas em grupo ou a escrita de uma redação. Esse é o campo observacional, onde elementos da realidade são avaliados com base em informações mais restritas e mensuráveis.

Por fim, temos o campo da decisão, onde, com base no conjunto limitado de informações que se imagina que medem algum aspecto da realidade, nós tomamos a decisão. O campo da decisão pode ser muito estreito. No nosso caso, elas se resumem a opções binárias do tipo *sim* ou *não*, como aceitar ou não aceitar determinada candidata para uma vaga.

O funcionamento de decisões algorítmicas trabalha da seguinte maneira: tenta transformar um conjunto de características observáveis em um punhado de informações, e, a partir disso, se faz uma previsão sobre a realidade. A decisão algorítmica percorre então essas transformações entre campos, saindo de uma realidade infinitamente complexa e culminando em uma decisão. A cada salto, restringe-se o campo analisável a partir de uma série de vieses e a decisão final jamais terá acesso a toda riqueza de informação da realidade ou mesmo dos campos anteriores.

Por conta disso, a própria natureza do funcionamento algorítmico força vieses e o reducionismo. Esses vieses podem ser estruturais, como na própria decisão do construto ou características a serem observadas.

Talvez a analogia mais evidente para se compreender esse ponto seja o vestibular, usado para selecionar alunos a uma universidade. Um dos vieses é a própria forma como a prova é conduzida, que ignora capacidades como inteligência emocional ou habilidade para trabalhar em grupo. Em um plano estrutural, também é sabido que o vestibular obedece a vieses estruturais, favorecendo classes sociais mais ricas, que têm meios para pagar por cursos preparatórios. Os problemas com decisões algorítmicas são de natureza semelhante, como se para cada decisão ocorresse um processo de vestibular automatizado.

Nesse sentido, é nítido que o próprio processo de mensuração da realidade e tomada de decisão é inerentemente enviesado. Em cima desses problemas, colocamos que as “visões de mundo” do que pode ser justo são necessariamente

incompatíveis. Soluções utilitárias podem propor resultados diferentes de soluções igualitárias ou mesmo equitárias. As decisões necessariamente terão de se orientar por princípios conflitantes e de alguma forma tentar conciliá-los.

A razão disso remonta ao pensamento de Rawls, que identificava a justiça como “uma virtude das instituições sociais”[191]. Neste sentido, o que se deve esperar de instituições verdadeiramente justas é a ausência de arbitrariedades entre indivíduos, devendo existir um equilíbrio social garantido pela forma como tais instituições atribuem direitos (vantagens) e deveres (desvantagens).

Ou seja, ainda no paralelo com a regulação de leitos, percebe-se que a injustiça, segundo Rawls, não emana na escolha de salvar a vida de um paciente em detrimento de outro, mas da arbitrariedade e da personalização dessa decisão. A institucionalidade vem como uma ferramenta a favor da justiça, que operacionaliza decisões necessárias de alocação de vantagens e desvantagens, mas conforme uma parametrização essencialmente normativa. Nesse sentido, a instituição, entendida aqui como conjunto de regras formais e informais que baliza o comportamento humano, é o recurso que mitiga a arbitrariedade casuística humana nessas decisões. O casuísmo é o elemento que gera incerteza e descrédito nas relações dos indivíduos com sistemas que geram alocações de recursos na sociedade.

O que se percebe é que este equilíbrio não significa estabelecer exatamente os mesmos deveres e direitos a todos. Na verdade, é justamente o oposto disso: é a certeza de que em cada circunstância haverá uma alocação de vantagens e de desvantagens correspondente. O que permanece inalterada é a institucionalidade, que será o conjunto de regras que assegura o sopesamento justo de todos os valores relevantes na decisão que resultará na distribuição de direitos.

Por isso, a concepção de justiça como equidade se constitui a partir de princípios de justiça, que têm a função de auxiliar as instituições a identificarem quais são as características de cada indivíduo que devem ser consideradas para que a distribuição de direitos e deveres, ainda que desigual, permita o equilíbrio de vantagens sociais. Dessa forma, é essencial ter bem definidos os princípios de justiça, sendo que a falta de um parâmetro pode fazer com que as instituições não pareçam justas para os demais indivíduos.

Aqui cabe relembrar: algoritmos são procedimentos. Em um contexto em que

---

[191] RAWLS, J. *Uma Teoria da Justiça*. 3. ed. São Paulo: Martins Fontes, 2008.

infraestruturas públicas e privadas são essencialmente geridas por sistemas da informação e comunicações, os algoritmos são tanto o ferramental de automação que dão escalabilidade ao trabalho humano, como se tornam os procedimentos que concretizam alocações de recursos na sociedade. A compreensão dos procedimentos e o seu questionamento são peças essenciais desse processo.

Assim, eles executam em essência um papel de alocação de direitos, ainda que isso se opere com fulcro em um instrumento jurídico privado, como os termos de uso de uma plataforma, ou com base em normas jurídicas de mais alta hierarquia, como o algoritmo do auxílio emergencial que age na execução das várias portarias e leis que estabeleceram a distribuição de recursos para a sociedade brasileira. O dever de explicação surge como uma obrigação transversal, um componente obrigatório da existência de procedimentos que concretizam alocações de vantagens, visto que junto ao direito de se conhecer as regras dessas institucionalidades, cabe, como corolário, o direito de entender como a institucionalidade operou caso a caso e de questioná-la, se necessário.

As instituições consideradas não justas geram desconfiança e passam a ser descredibilizadas. Por essa razão também é necessário que tais instituições sigam os princípios de justiça definidos pelos indivíduos, resultado de um consenso entre os mesmos. A única forma de sustentar o crédito é através da publicidade não só das decisões, mas do processo de decisão — o direito à explicação — algo que não é nada estranho ao mundo jurídico, se comparado ao dever de publicidade de decisões judiciais e administrativas, por exemplo, pilares do devido processo legal, que será visto a seguir.

Contudo, aqui tratamos de um dever mais amplo que transcende uma relação de transparência da administração pública. A explicação surge como um mecanismo para assegurar a simetria nas relações em que uma pessoa vê seus interesses sujeitos a decisões executadas pela tecnologia.

2.2.3 Direito à explicação como corolário do devido processo informacional

O texto “Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms”[192], de autoria de Kate Crawford e Jason Schultz, apresenta os

[192] CRAWFORD, K.; SCHULTZ, J. *Big Data and Due Process:* Toward a Framework to Redress

desafios, para os defensores da privacidade, trazidos pelo advento da tecnologia de *Big Data* no setor privado e pelo fato de sua configuração muitas vezes “fugirem” das regulações tradicionais, criando um modelo de tratamento de dados que, ainda que eventualmente não identifique indivíduos específicos, têm efeitos sobre suas vidas. O artigo propõe uma nova abordagem para endereçar e mitigar riscos à privacidade nesse contexto — a de um direito ao “devido processo de dados”, partindo de uma análise do papel do devido processo legal no sistema anglo-americano.

A abordagem é semelhante à de Hildebrandt e Citron[193] e discorre sobre as diferenças do modelo de *profiling* automatizado atual. O artigo aponta uma mudança importante trazida pelo *Big Data*, que é a ampliação do escopo de dados que podem ser considerados identificáveis. Com base em informações publicamente disponíveis, processos de *Big Data* podem gerar um modelo preditivo que essencialmente *imagina* os dados de um indivíduo e é potencialmente tão sensível quanto outros usos, sem, no entanto, submeter-se a regulações tradicionais (à época, ao menos).

Dessa forma, não se trata apenas de coleta e análise, mas efetivamente de geração de novos dados, o que impõe um desafio ao sistema estadunidense no endereçamento de riscos à privacidade, que tradicionalmente divide-se nos momentos de coleta, tratamento e divulgação de dados. Por ser uma ferramenta de correlação, e não causalidade, a própria lógica de *analytics* do *Big Data* pode escapar a regulações tradicionais de privacidade e proteção de dados. Isso justifica a necessidade, segundo Crawford e Schultz, de um novo devido processo legal aplicado a esta dinâmica.

Historicamente, a construção doutrinária e jurisprudencial do devido processo legal nos Estados Unidos caracteriza-o como uma “[...] vedação à privação de direitos como a vida, a integridade e a liberdade”[194]. Crawford e Schultz argumentam que esta mesma lógica pode se aplicar à privacidade e proteção de dados pessoais pelo prisma das liberdades positivas e negativas e que se trata de uma abordagem preferível ao foco exclusivo no arcabouço jurídico de proteção de dados estadunidense consolidado no que ficou conhecido como os *Fair Information Practices* (FIPs). A abordagem no

---

Predictive Privacy Harms. Rochester, NY: Social Science Research Network, 2013. Disponível em: https://papers.ssrn.com/abstract=2325784. Acesso em: 27 maio 2020.

[193] CITRON, D. K. Technological Due Process*. Washington University Law Review*, v. 85, p. 1249-1313, 2007. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1012360. Acesso em: 21 jun. 2021.

[194] ESTADOS UNIDOS DA AMÉRICA. Constituição dos Estados Unidos da América. 1787. Emendas VI e XIV.

devido processo absorve a lógica da adjudicação, focando precisamente no direito de "auditar" o dado pessoal usado para fazer uma inferência ou decisão específica. O mesmo se aplica em relação às obrigações de notificação — que no caso do devido processo seriam específicas quanto à ação que afeta o indivíduo e não apenas sobre a coleta ou tratamento de dados de forma genérica.

Idêntico esforço de justificação da necessidade de reconhecimento de um devido processo informacional no contexto norte-americano foi realizado também por Citron e Pasquale, conforme pontuado anteriormente. Em 2014, os autores escreveram o *paper* "The Scored Society: Due Process for Automated Predictions", que avança nas reflexões apresentadas já em 2007 por Citron[195], no qual analisam a crescente presença de decisões automatizadas no nosso cotidiano, frequentemente cercadas de opacidade e com impactos significativos na esfera de direitos dos indivíduos, em especial de grupos historicamente marginalizados e estigmatizados.

Os autores reconhecem que sistemas automatizados, como os de pontuação de crédito, vieram para ficar e fazem parte da atual Era Informacional, mas que o seu funcionamento não pode permanecer sem escrutínio. Neste sentido, argumentam que é necessário cercar esses sistemas de certa regularidade procedimental, uma vez que que eles atuam diretamente ampliando ou limitando as oportunidades de vida dos indivíduos nos mais variados contextos. Nesse sentido, afirmam: "*When scoring systems have the potential to take a life of their own*, contributing to or creating the situation they claim merely to predict, it becomes a normative *matter, requiring moral justification and rationale*."[196] Seria, portanto, necessário sujeitar os sistemas automatizados a alguns requisitos de justiça que reflitam o impacto proporcional desses mecanismos na vida dos indivíduos.[197]

A forma proposta pelos autores para cercar esses algoritmos de *accountability* consiste em estabelecer salvaguardas que podem ser reunidas em torno da ideia de "[...] technological due process — procedures ensuring that predictive algorithms live up to some standard of review and revision to ensure their fairness and accuracy."[198]

---

[195] CITRON, D. K. Technological Due Process*. Washington University Law Review*, v. 85, p. 1249-1313, 2007. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1012360. Acesso em: 21 jun. 2021.
[196] Ibidem, p. 18.
[197] CITRON, D. K.; PASQUALE, F. The Scored Society: Due Process for Automated Predictions. *Washington Law Review*, v. 89, 2014. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2376209. Acesso em: 2 abr. 2021. p. 19.
[198] Ibidem, p…

Os autores propõem um modelo de devido processo informacional de escopo amplo, aplicável não apenas ao setor público, mas também a atores privados:

> This is not to suggest that full due process guarantees are required as a matter of current law. Given the etiolated state of "state action" doctrine in the United States, FICO and credit bureaus are not state actors; however, much of their business's viability depends on the complex web of state supports and rules surrounding housing finance. Nonetheless, the underlying values of due process—transparency, accuracy, accountability, participation, and fairness—should animate the oversight of scoring systems given their profound impact on people's lives. Scholars have built on the "technological due process" model to address private and public decision-making about individuals based on the mining of Big Data.[199]

Os autores propõem um modelo de devido processo informacional baseado em dois elementos: (i) controle regulatório sobre sistemas algorítmicos; e (ii) garantia de um conjunto de salvaguardas aos indivíduos. Embora os autores tenham utilizado os sistemas de *credit scoring* como referência e ponto de partida para a elaboração de suas recomendações, eles argumentam que o modelo de devido processo informacional proposto podem ser estendido a outros tipos de algoritmos preditivos que tenham o potencial de gerar impactos negativos na vida dos indivíduos.[200]

O primeiro elemento, portanto, consiste em cercar os sistemas algorítmicos de algum grau de controle regulatório. Nesse sentido, Citron e Pasquale pontuam que um primeiro passo consiste em identificar as etapas existentes no processo de *scoring*, quais sejam: (i) coleta e combinação de dados; (ii) transformação dos dados em um *score*; (iii) divulgação do *score* a *decision makers* (em sentido amplo); e (iv) uso dos *scores* em processos de tomadas de decisão.

Na primeira etapa, argumentam: "*Individuals should have the right to inspect, correct, and dispute inaccurate data, and to know the sources (furnishers) of the data*."[201] Na segunda fase, deve se garantir algum nível de transparência e escrutínio público em torno da forma como os dados são valorados e processados:

> Second, at the calculation of data stage, ideally such calculations would be public, and all processes (whether driven by AI or other

---

[199] CITRON, D. K.; PASQUALE, F. The Scored Society: Due Process for Automated Predictions. *Washington Law Review*, v. 89, 2014. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2376209. Acesso em: 2 abr. 2021. p. 19-20.
[200] Ibidem, p. 20.
[201] CITRON, D. K.; PASQUALE, F. The Scored Society: Due Process for Automated Predictions. *Washington Law Review*, v. 89, 2014. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2376209. Acesso em: 2 abr. 2021.

> computing) would be inspectable. In some cases, the trade secrets may merit protection, and only a dedicated, closed review should be available. But in general, we need to switch the default in situations like this away from an assumption of secrecy, and toward the expectation that people deserve to know how they are rated and ranked.[202]

Na terceira fase, deve-se garantir que o indivíduo tenha conhecimento sobre o fluxo de seus dados, notificando-o quando sua pontuação for comunicada a determinada entidade: "Nevertheless, scored individuals should be notified when scores or data are communicated to an entity."[203]

A quarta e última fase, de acordo com os autores, pode ser considerada como a mais controversa de todas. Nesse estágio, quando os scores são utilizados para a tomada de decisão, os autores argumentam que deve haver algum tipo de licenciamento ou auditoria dos sistemas automatizados envolvidos, sobretudo quando se tratar de sistemas automatizados que possam desencadear um impacto sobre grupos vulneráveis, como é o caso de sistemas automatizados mediando relações de emprego, seguros e serviços de saúde, por exemplo.[204]

Mas como operacionalizar essas demandas por maior transparência em torno de sistemas automatizados quando nem mesmo os reguladores estão aptos a compreender inteiramente como essas *black boxes* funcionam? Neste ponto, os autores sugerem duas soluções: (i) garantir maior acesso, transparência e abertura desses sistemas opacos com potencial de lesar os consumidores aos órgãos reguladores; e (ii) elaboração de relatórios de análise de risco e recomendações sobre o funcionamento desses sistemas, com abertura dos dados ao público amplo, ou seja, não apenas aos titulares de dados, mas também a acadêmicos, especialistas e outros atores.

O segundo elemento do modelo de devido processo informacional proposto por Citron e Pasquale consiste em garantir salvaguardas aos indivíduos na forma de notificações:

> In constructing strategies for technological due process in scoring contexts, it is helpful to consider the sort of notice individuals are owed when governmental systems make adverse decisions about them. Under the Due Process Clause, notice must be "reasonably calculated"

---

[202] Ibidem, p. 21.
[203] CITRON, D. K.; PASQUALE, F. The Scored Society: Due Process for Automated Predictions. *Washington Law Review*, v. 89, 2014. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2376209. Acesso em: 2 abr. 2021.
[204] Ibidem, p. 21-22.

> to inform individuals of the government's claims against them. The sufficiency of notice depends upon its ability to inform affected individuals about the issues to be decided, the evidence supporting the government's position, and the agency's decisional process. Clear notice decreases the likelihood that agency action will rest upon "incorrect or misleading factual premises or on the misapplication of rules.[205]

Importa salientar que essas notificações deverão ser completas, adequadas, úteis aos indivíduos para fins de exercício dos seus direitos. Os autores argumentam que notificações desse tipo podem ser garantidas se existirem trilhas de auditoria capazes de registrar o funcionamento dos algoritmos:

> Aggrieved consumers could be guaranteed reasonable notice if scoring systems included audit trails recording the correlations and inferences made algorithmically in the prediction process. With audit trails, individuals would have the means to understand their scores. They could challenge mischaracterizations and erroneous inferences that led to their scores. Even if scorers successfully press to maintain the confidentiality of their proprietary code and algorithms vis-à-vis the public at large, it is still possible for independent third parties to review it. One possibility is that in any individual adjudication, the technical aspects of the system could be covered by a protected order requiring their confidentiality. Another possibility is to limit disclosure of the scoring system to trusted neutral experts. Those experts could be entrusted to assess the inferences and correlations contained in the audit trails. They could assess if scores are based on illegitimate characteristics such as race, nationality, or gender or on mischaracterizations. This possibility would both protect scorers' intellectual property and individuals' interests.[206]

Além de garantir trilhas de auditoria dos algoritmos, uma outra forma de cercar esses sistemas de transparência seria o que os autores chamam de *interactive modelling*, mecanismo bastante similar à ideia de *conterfactual explanations* proposta por Wachter *et al.*: "Another approach would be to give consumers the chance to see what happens to their score with different hypothetical alterations of their credit histories."[207]

Garantir um maior nível de transparência em torno desses sistemas não é tarefa fácil, contudo, sobretudo em razão da grande objeção apresentada por algumas organizações. De acordo com os autores, os *bureaus* de crédito, por exemplo, tendem

---

[205] Ibidem, p. 27.
[206] CITRON, D. K.; PASQUALE, F. The Scored Society: Due Process for Automated Predictions. *Washington Law Review*, v. 89, 2014. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2376209. Acesso em: 2 abr. 2021. p. 28.
[207] Ibidem, p. 28-29.

a apresentar objeções a demandas por transparência em torno do funcionamento de seus sistemas de *credit scoring*. Disponibilizar essas informações ao público, argumentam, colocaria em xeque a própria lógica da análise de crédito, além de abrir margem para que os indivíduos manipulem o sistema (*game the system*) ou incorram em comportamento fraudulento. Os autores fazem uma concessão a esse argumento, mas se contrapõem afirmando que a existência de sistemas opacos decidindo sobre aspectos relevantes da vida das pessoas não é algo que deve se sobrepor à demanda por transparência.[208]

É com base nesta construção teórica em torno da ideia de devido processo informacional, derivada da cláusula geral do devido processo do sistema constitucional norte-americano, que argumentamos inicialmente ser possível reconhecer a existência de um direito à explicação nos EUA, mesmo na ausência de leis expressamente prevendo este direito. O reconhecimento do direito à explicação por meio do devido processo informacional soma-se, portanto, às outras formas anteriormente apresentadas de justificação deste direito, que coexistem como fundamentos ao lado dos elementos disponíveis no direito positivado, conforme passaremos a expor no capítulo a seguir.

Por fim, cabe salientar que a noção de devido processo informacional, apesar de ter sido gestada no cenário anglo-saxão, começa a ser gradativamente incorporada à jurisprudência constitucional brasileira. O conceito foi invocado, como vimos no tópico 2.2.1.3, no julgamento da ADI nº 6.387[209] e, mais recentemente, no contexto do julgamento da ADI nº 6.649/DF, que analisa a constitucionalidade do Cadastro Base do Cidadão, instituído pelo Decreto nº 10.046/2019:

> O regime de proteção de dados pessoais é, portanto, eminentemente procedimental: objetiva não a vedação do tratamento, mas sim sua efetivação de forma adequada. Assim, o devido processo informacional apresenta-se como um conceito fundamental para o caso em questão, garantindo que os dados possam estar em fluxo de

[208] CITRON, D. K.; PASQUALE, F. The Scored Society: Due Process for Automated Predictions. *Washington Law Review*, v. 89, 2014. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2376209. Acesso em: 2 abr. 2021. p. 30-32.

[209] "A partir da tradição norte-americana, também é possível identificar como corolário da dimensão subjetiva do direito à proteção de dados pessoais, a preservação de verdadeiro "devido processo informacional" (*informational due process privacy right*), voltado a conferir ao indivíduo o direito de evitar exposições de seus dados sem possibilidades mínimas de controle, sobretudo em relação a práticas de tratamento de dados capazes de sujeitar o indivíduo a julgamentos preditivos e peremptórios." Cf. BRASIL. Supremo Tribunal Federal. *Medida Cautelar na Ação Direta de Inconstitucionalidade nº 6.387 Distrito Federal*. Relator: Min. Rosa Weber. Data de Julgamento: 24/04/2020. Disponível em: http://www.stf.jus.br/arquivo/cms/noticiaNoticiaStf/anexo/ADI6387MC.pdf. Acesso em: 25 nov. 2020.

> forma justa. Por devido processo informacional entende-se não só assegurar direitos aos titulares de contestação acerca do uso de seus dados, mas também a fixação de deveres por parte dos agentes de tratamento de dados para que a sua interferência seja justa. Trata-se tanto de um instrumento de garantias processuais em sede judicial, quanto de uma ferramenta para assegurar a simetria e proporcionalidade de forma mais ampla em relações Estado-indivíduo e privadas, no que tange ao tratamento de dados pessoais. O devido processo informacional materializa a exigência do regime de proteção de dados pessoais no sentido de garantir o controle sobre dados pessoais, no sentido de estabelecer mecanismos (direitos e obrigações) de combate e mitigação dos riscos ao titular decorrentes do tratamento de seus dados. Apesar de soar conceito novíssimo, trata-se de um conjunto de noções jurídicas sobre garantias fundamentais profundamente enraizadas nas tradições jurídicas americanas, do Norte e do Sul.[210]

2.3 SÍNTESE DO CAPÍTULO

Neste capítulo debatemos como a existência de um direito à explicação no contexto europeu esteve muito marcado pelas possíveis interpretações dos termos da GDPR. Neste sentido, diante da ausência de previsão expressa, a resposta depende de uma análise mais profunda sobre os fundamentos jurídicos da existência de um direito. Apontamos como alguns autores, a partir de uma interpretação da literalidade da lei, entendem que o sistema jurídico resultante do GDPR não permite a afirmação de que existe um direito à explicação. Enquanto outros autores, a partir de uma interpretação sistemática, defendem a sua existência em função do regime de accountability consubstanciado a partir da aprovação do GDPR, dos *recitals* e das opiniões do WP 29 e o EDPB.

A partir dessa compreensão do direito a explicação inserido num regime de accountability, apresentamos o que seriam decisões automatizadas e o que seria uma explicação. Esse direito deve então ser compreendido como necessário frente a limitação do conceito de transparência para a garantia da proteção dos titulares de dados, visto que o mero dever de informação e de acesso delega aos titulares toda a responsabilidade pela sua proteção, desconsiderando as assimetrias informacionais. Além disso, o excesso de transparência pode gerar um efeito inverso, de diminuir as

[210] SECAF, H.; ZANATTA, R. A. F.; NUÑEZ, I. S. O Cadastro Base do Cidadão na mira do Supremo. *Jota*, 9 abr. 2021. Disponível em: https://www.jota.info/opiniao-e-analise/colunas/agenda-da-privacidade-e-da-protecao-de-dados/o-cadastro-base-do-cidadao-na-mira-do-supremo-09042021. Acesso em: 24 abr. 2021.

possibilidades de ação e controle do titular de dados, incapaz de compreender a complexidade do tratamento de dados ao qual estaria sujeito.

É a partir dessa constatação que apresentamos os fundamentos jurídicos do direito à explicação como um elemento essencial à proteção dos direitos da personalidade. Essa fundamentação partiu da compreensão de que o sistema jurídico deve ser unitário, coerente e completo, de forma que é possível afirmar a existência de um direito a partir da integração de elementos internos ou externos ao sistema jurídico, para preencher lacunas legislativas ou doutrinárias.

Tendo em vista os precedentes no campo do direito consumerista, o reconhecimento da autodeterminação informativa como direito fundamental, as previsões de transparência presentes na regulação de proteção de dados e a noção de que decisões automatizadas podem violar a dignidade da pessoa humana, o direito à explicação se apresenta como uma salvaguarda para os direitos da personalidade em contextos de assimetrias informacionais e do tratamento automatizado de dados capaz de afetar direitos e determinar o acesso a bens e serviços.

Por fim, este capítulo apresentou o conceito de devido processo informacional, da forma como aparece no contexto estadunidense, onde a questão sobre a governança e a transparência algorítmica também ocorria em torno de seu ordenamento jurídico. Mencionado no voto do ministro Gilmar Mendes na ADI 683, tendo em vista a relevância que o conceito de devido processo encontra também no ordenamento brasileiro, pode-se considerar a noção como um elemento em favor da existência do direito à explicação.

No próximo capítulo, após essa fundamentação doutrinária, discutiremos mais detalhadamente os diferentes textos normativos e as suas implicações para o direito à explicação nos diferentes ordenamentos analisados neste trabalho: o ordenamento estadunidense, o regulamento europeu e o plexo normativo brasileiro, discutindo como o direito a explicação se insere no campo da proteção de dados e em outras disciplinas do direito.

3 ELEMENTOS REGULATÓRIOS PARA O RECONHECIMENTO DO DIREITO À EXPLICAÇÃO NO CONTEXTO DE DECISÕES AUTOMATIZADAS

No Capítulo 2, discutimos os possíveis fundamentos metajurídicos para o reconhecimento de um direito à explicação, abordando formas de compreensão de um direito que vão além daquelas expressamente previstas na legislação e regulamentos. Nesse sentido, buscamos desenvolver a fundamentação do direito à explicação enquanto um direito da personalidade, derivado da autodeterminação informativa e associado à dignidade da pessoa humana. Em seguida, procuramos demonstrar como o direito à explicação pode ser derivado da necessidade de garantir que os sistemas automatizados operem de forma justa. Por fim, desenvolvemos o argumento no sentido de situar o direito à explicação como um corolário do devido processo informacional, princípio derivado da cláusula geral do devido processo legal.

No presente tópico, apresentaremos os elementos regulatórios para o reconhecimento do direito à explicação no contexto de decisões automatizadas, tomando como referência três jurisdições: o cenário norte-americano, o cenário europeu e o cenário brasileiro. Em cada ordenamento, buscamos identificar em que medida a legislação e a jurisprudência oferecem elementos aptos a fundamentar a existência de um direito à explicação.

3.1 CENÁRIO REGULATÓRIO INTERNACIONAL

Na ausência de uma previsão legal expressa, a incerteza em torno da (in)existência de um direito à explicação tem levado inúmeros autores, em diferentes jurisdições e a partir de elementos e configurações normativas distintos, a realizarem um exercício de fundamentação deste direito a partir dos instrumentos jurídicos atualmente existentes.

No cenário europeu, a discussão tem se valido de elementos presentes na própria GDPR, bem como da atividade interpretativa do Article 29 Working Party, do European Data Protection Board, dos tribunais regionais e das autoridades nacionais de proteção de dados no âmbito de cada Estado-Membro, concentrando-se sobretudo em torno das disposições que tratam sobre obrigações de transparência, direitos de acesso e das salvaguardas previstas no contexto de decisões automatizadas, principalmente em suas lógicas subjacentes.

Nos Estados Unidos, apesar de não haver uma legislação geral de proteção de dados pessoais, muito menos uma autoridade especializada de *enforcement* ou disposições específicas versando sobre o direito à explicação, este tem sido derivado a partir da cláusula geral do processo legal, conforme previsto na 5ª e 14ª emendas da Constituição dos Estados Unidos, que, uma vez contextualizada, dá origem à noção de devido processo informacional, construção que dá abertura ao reconhecimento de um direito à explicação no contexto norte-americano.

Tendo em vista a centralidade desses dois cenários no atual debate sobre o direito à explicação, é bastante provável que eles venham a influenciar o debate brasileiro sobre a existência de um direito à explicação na LGPD. Ademais, o mapeamento do debate internacional nos permite aprender com os erros e acertos da experiência jurídica internacional, bem como ter contato com elementos e soluções que podem ser transpostos para o cenário brasileiro, observadas suas particularidades e especificidades.

3.1.1 Estados Unidos

Ao analisar a regulação da proteção de dados pessoais ao redor do mundo, Abraham L. Newman[211] categoriza as diferentes formas de regulação existentes em dois regimes. O primeiro deles, identificado como regime de proteção de dados pessoais compreensivo, pode ser caracterizado, dentre outros aspectos, pela regulação de escopo amplo, geral, estendendo-se tanto sobre atores públicos quanto sobre atores privados, pela existência de uma regulação supervisionada por uma autoridade central de *enforcement* e por uma preocupação em balancear a proteção aos direitos do titular e a dimensão econômica dos dados. O segundo, referido como regime de proteção de dados pessoais de escopo limitado, é caracterizado pela ausência de uma regulação geral e compreensiva sobre a matéria, optando, por sua vez, por privilegiar a autorregulação e se valer de mecanismos de mercado para a correção de eventuais abusos no tratamento de dados pessoais. Nesse segundo modelo, a regulação estatal ocorre de forma pontual e isolada, alcançando apenas alguns setores específicos e de maior sensibilidade. É comum, ainda, observar a ausência de uma autoridade regulatória central responsável por supervisionar as

[211] NEWMAN, A. P. *Protectors of Privacy*: Regulating Personal Data in the Global Economy. London: Cornell University Press, 2008. p. 26-35.

atividades de *enforcement* nos países que adotam esse regime.

Quando analisado, o regime de proteção de dados pessoais existente nos Estados Unidos parece se filiar ao modelo de regulação de escopo limitado.[212] Atualmente, os EUA não contam com uma lei geral de proteção de dados pessoais de caráter compreensivo e abrangente a nível federal, dispondo, contudo, de um mosaico de normativas setoriais endereçando o tratamento de dados pessoais, que incluem, dentre outras legislações, o *Electronic Communications Privacy Act*[213], de 1986, composto pelo *Wiretap Act*, o *Stored Communications Act* e o *Pen Register Act*, responsável por regular o tratamento de dados pessoais em fluxo ou armazenados no contexto de comunicações telemáticas, o *Children's Online Privacy Protection Act* (COPPA)[214], de 1998, que traz disposições sobre proteção de dados pessoais e da privacidade de crianças abaixo de 13 anos no ambiente *online*, o *Health Insurance Portability and Accountability Act* (HIPAA)[215], de 1996, que regula aspectos de privacidade e proteção de dados no setor de saúde, e o *Privacy Act*[216], de 1974, que regula o tratamento de dados pessoais no contexto de atividades públicas conduzidas pelas agências federais.[217] Conta, ainda, com normas estaduais, em destaque para a *California Consumer Privacy Act* (CCPA), de 2018.

Nesse contexto de descentralização normativa, há também uma miríade de órgãos encarregados pelo *enforcement* de cada instrumento normativo em seu respectivo âmbito setorial:

> O sistema estadunidense não possui uma Autoridade de Proteção de Dados nos moldes europeus, um órgão técnico, independente e dedicado unicamente à matéria da privacidade e da proteção de dados pessoais. Em seu lugar, certos órgãos já existentes e não exclusivos do governo atuam como agências reguladoras, sendo responsáveis

---

[212] "Regimes governing the processing of personal data can be broadly categorized in two ways. The first category is omnibus; the EU regime is categorized in this way. The second category of regime is sectoral, also sometimes referred to as a 'sectional' or 'limited' regime. The regime in place in the United States (US) typifies, albeit perhaps to a decreasing extent, this type of 'sectoral' regime." (LYNSKEY, O. *The Foundations of EU Data Protection Law*. Oxford: Oxford University Press, 2015. p. 15).

[213] ESTADOS UNIDOS DA AMÉRICA. *Electronic Communications Privacy Act, 18 U.S.C. §2510 e ss., Public Law*. Washington D.C., 21 out. 1986.

[214] Idem. *Children's Online Privacy Protection Act, 15 U.S.C*. §6501-6506., Public Law, Washington D.C., 21 out. 1998.

[215] Idem. *Health Insurance Portability and Accountability Act*. 110 Stat. 1936, Public Law, Washington D.C., 21 ago. 1996.

[216] Idem. *Privacy Act, 88 Stat. 1896 Public Law*, Washington D.C., 31 dez. 1974.

[217] Para uma análise do regime de proteção de dados pessoais norte-americano, cf. GUIDI, Guilherme Berti de Campos. Modelos regulatórios para proteção de dados pessoais. Rio de Janeiro: Instituto de Tecnologia e Sociedade do Rio, 2017. Disponível em: https://itsrio.org/wp-content/uploads/2017/03/Guilherme-Guidi-V-revisado.pdf. Acesso em: 22 mar. 2021. p. 10-15.

> pelo enforcement das leis vigentes, igualmente separados por setores econômicos, de modo que sua atuação não é necessariamente homogênea, como também não são necessariamente suas posições em relação às controvérsias que possam surgir sobre este ou aquele conceito. Assim, a Federal Trade Commission é responsável, por exemplo, por fiscalizar a aplicação do COPPA e das regras relativas à proteção do consumidor, segundo seu próprio estatuto, que podem incluir abusos na coleta e utilização de dados dos consumidores. Já o Department of Health and Human Services, é responsável pela supervisão do cumprimento do HIPAA. No setor financeiro, temos o Consumer Financial Protection Bureau.[218]

Dentre as várias legislações existentes no contexto norte-americano que abordam, direta ou indiretamente, privacidade e proteção de dados pessoais, não há nenhuma endereçando especificamente o direito à explicação no contexto de decisões automatizadas. O que há de mais próximo de uma regulação nesse sentido são as disposições contidas na Seção 5(a) do *Federal Trade Commission Act* (FTC Act), que tratam sobre os *Deceptive Acts or Practices*. Nos termos da *FTC Policy Statement on Deception*, documento que busca traçar diretrizes sobre a aplicação e interpretação do disposto na Seção 5(a) do FTC Act, *deceptive acts or practices* podem ser entendidos como:

> "Deceptive" practices are defined [...] as involving a material representation, omission or practice that is likely to mislead a consumer acting reasonably in the circumstances. An act or practice is "unfair" if it "causes or is likely to cause substantial injury to consumers which is not reasonably avoidable by consumers themselves and not outweighed by countervailing benefits to consumers or to competition."[219]

A venda ou uso de um sistema automatizado dotado de certos vieses discriminatórios, por exemplo, pode cair no escopo da Seção 5 do FTC Act, conforme pontuado pela Federal Trade Commission no *press release* "Aiming for truth, fairness, and equity in your company's use of AI", em abril de 2021. No mesmo documento, a agência destaca a possibilidade de exercer sua competência em casos envolvendo o

---

[218] GUIDI, Guilherme Berti de Campos. Modelos regulatórios para proteção de dados pessoais. Rio de Janeiro: Instituto de Tecnologia e Sociedade do Rio, 2017. Disponível em: https://itsrio.org/wp-content/uploads/2017/03/Guilherme-Guidi-V-revisado.pdf. Acesso em: 22 mar. 2021. p. 13.

[219] FEDERAL TRADE COMMISSION. A Brief Overview of the Federal Trade Commission's Investigative, Law Enforcement, and Rulemaking Authority. *FTO*, out. 2019. Disponível em: https://www.ftc.gov/about-ftc/what-we-do/enforcement-authority. Acesso em: 22 mar. 2021; Neste sentido, cf. também: FEDERAL TRADE COMMISSION. FTC Policy Statement on Deception. *FTC*, Washington, D.C., 14 out. 1983. Disponível em: https://www.ftc.gov/system/files/documents/public_statements/410531/831014deceptionstmt.pdf. Acesso em: 22 mar. 2021.

uso de sistemas automatizados com base no “Fair Credit Reporting Act” e no “Equal Credit Opportunity Act”:

> The FCRA comes into play in certain circumstances where an algorithm is used to deny people employment, housing, credit, insurance, or other benefits. The ECOA makes it illegal for a company to use a biased algorithm that results in credit discrimination on the basis of race, color, religion, national origin, sex, marital status, age, or because a person receives public assistance.[220]

Como visto acima, não há, no cenário norte-americano, uma lei geral de proteção de dados pessoais nem mesmo qualquer normativa setorial tratando especificamente sobre o direito à explicação. Não obstante, conforme buscaremos desenvolver ao longo deste capítulo, mesmo sem uma lei geral de proteção de dados pessoais, e na ausência de previsões normativas específicas, há um sistema que, partindo das obrigações de transparência e da cláusula geral do devido processo legal do constitucionalismo norte-americano, nos permite trabalhar com a ideia de um direito à explicação em sistemas automatizados no seu cenário regulatório. Nesse sentido, passaremos a apresentar a seguir os fundamentos da ideia de um devido processo informacional nos EUA, quais sejam, os *Fair Information Practice Principles* (FIPPs), bem como o debate teórico desenvolvido em torno desses princípios e do direito à explicação sob a ótica do devido processo informacional.

O texto “Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms”[221], de autoria de Kate Crawford e Jason Schultz, apresenta os desafios, para os defensores da privacidade, trazidos pelo advento da tecnologia de *Big Data* no setor privado e pelo fato de sua configuração muitas vezes “fugir” das regulações tradicionais, criando um modelo de tratamento de dados que, ainda que eventualmente não identifique indivíduos específicos, têm efeitos sobre suas vidas. O artigo propõe uma nova abordagem para endereçar e mitigar riscos à privacidade nesse contexto — a de um direito ao “devido processo de dados”, partindo de uma análise do papel do devido processo legal no sistema anglo-americano.

---

[220] JILLSON, E. Aiming for truth, fairness, and equity in your company’s use of AI. *Federal Trade Commision*, 19 abr. 2021. Disponível em: https://www.ftc.gov/news-events/blogs/business-blog/2021/04/aiming-truth-fairness-equity-your-companys-use-ai. Acesso em: 22 jun. 2021.

[221] CRAWFORD, K.; SCHULTZ, J. *Big Data and Due Process:* Toward a Framework to Redress Predictive Privacy Harms. Rochester, NY: Social Science Research Network, 2013. Disponível em: https://papers.ssrn.com/abstract=2325784. Acesso em: 27 maio 2020.

A abordagem é semelhante à de Hildebrandt e Citron[222] e discorre sobre as diferenças do modelo de *profiling* automatizado atual. O artigo aponta uma mudança importante trazida pelo *Big Data*, que é a ampliação do escopo de dados que podem ser considerados identificáveis. Com base em informações publicamente disponíveis, processos de *Big Data* podem gerar um modelo preditivo que essencialmente *imagina* os dados de um indivíduo e é potencialmente tão sensível quanto outros usos, sem, no entanto, submeter-se a regulações tradicionais (à época, ao menos).

Dessa forma, não se trata apenas de coleta e análise, mas efetivamente de geração de novos dados, o que impõe um desafio ao sistema estadunidense no endereçamento de riscos à privacidade, que tradicionalmente divide-se nos momentos de coleta, tratamento e divulgação de dados. Por ser uma ferramenta de correlação, e não causalidade, a própria lógica de *analytics* do *Big Data* pode escapar a regulações tradicionais de privacidade e proteção de dados. Isso justifica a necessidade, segundo Crawford e Schultz, de um novo devido processo legal aplicado a esta dinâmica.

Historicamente, a construção doutrinária e jurisprudencial do devido processo legal nos Estados Unidos caracteriza-o como uma "[...] vedação à privação de direitos como a vida, a integridade e a liberdade."[223] Crawford e Schultz argumentam que esta mesma lógica pode se aplicar à privacidade e proteção de dados pessoais pelo prisma das liberdades positivas e negativas e que se trata de uma abordagem preferível ao foco exclusivo no arcabouço jurídico de proteção de dados estadunidense consolidado no que ficou conhecido como os *Fair Information Practices* (FIPs).

A abordagem no devido processo absorve a lógica da adjudicação, focando precisamente no direito de "auditar" o dado pessoal ou processo envolvido para fazer uma inferência ou decisão específica. O mesmo se aplica em relação às obrigações de notificação — que no caso do devido processo seriam específicas quanto à ação que afeta o indivíduo e não apenas sobre a coleta ou tratamento de dados de forma genérica.

A bibliografia sobre privacidade e proteção de dados destaca a importância da criação dos 5 princípios de *Fair Information Practices* (FIP), apresentados no relatório

---

[222] CITRON, D. K. Technological Due Process*. Washington University Law Review*, v. 85, p. 1249-1313, 2007. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1012360. Acesso em: 21 jun. 2021. p. 1249.

[223] ESTADOS UNIDOS DA AMÉRICA. Constituição dos Estados Unidos da América. 1787. Emendas V e XIV.

Records, Computers and the Rights of Citizens, de 1973. O processo de criação dos princípios presentes nesse relatório ocorreu no contexto das reuniões do *Secretary's Advisory Committee on Automated Personal Data Systems* (SACAPDS) do *Department of Health, Education and Welfare* (HEW).

O relatório apresentou, baseado na análise dos diversos usos de bancos de dados públicos e privados, os principais riscos na ótica de grandes especialistas da época. Além disso, apresentou recomendações de medidas de segurança e limites de utilização para garantir os direitos dos titulares. Os princípios apresentados são a base de uma agenda regulatória que representa o momento de emergência de novas tecnologias e configurações sociais, já premente à época.

O comitê foi composto por *experts*, profissionais de agências regulatórias, do poder legislativo e de setores econômicos. Foram nove encontros ocorridos entre 1972 e 1973. Uma análise do encontro e dos debates pode ser encontrada em artigo de Chris Jay Hoofnagle[224]. As questões debatidas são relevantes até os dias de hoje e permitem apreender o contexto do surgimento das preocupações sobre proteção de dados e privacidade que ocorreu em diversos países durante a década de 1970.

Esse contexto, que motivou a criação de diversos marcos legais para proteção de dados, é bem apresentado na obra de Colin Bennett[225]. Algumas características emergentes das sociedades pós-industriais levaram a convergências nas regulações. A rapidez da difusão da tecnologia da informação e sua opacidade; a criação de uma comunidade internacional em defesa da proteção de dados; uma tendência de padronização e cooperação internacional e a prevalência dos valores democráticos liberais fizeram com que, em diferentes países, os legisladores chegassem a soluções parecidas.

Embora os modelos de regulação variem de acordo com os arranjos institucionais de cada país, é comum indicar os FIPs como origem dos atuais instrumentos normativos de proteção de dados. São princípios cuja aplicação varia de acordo com os tipos de dados, a natureza do controlador e os objetivos do tratamento. Seu conteúdo pode ser encontrado em diversos instrumentos regulatórios dos EUA e

---

[224] HOOFNAGLE, C. J. *The Origin of Fair Information Practices: Archive of the Meetings of the Secretary's Advisory Committee on Automated Personal Data Systems* (SACAPDS). Berkeley Law, 16 jul. 2014. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2466418. Acesso em: 21 jun. 2021.

[225] BENNET, C. Regulating privacy: data protection and public policy in Europe and the United States. Ithaca, New York: Cornell University Press, 1992.

no mundo[226], como na GDPR e na LGPD.

Cinco princípios básicos, segundo o relatório, deveriam guiar a criação de um código nacional: a) não deve haver nenhum sistema de registro de dados pessoais cuja existência seja secreta; b) deve haver um meio para que os indivíduos descubram quais de seus dados pessoais são registrados e para que são usados; c) deve haver uma forma para os indivíduos evitarem que seus dados coletados para um propósito sejam utilizados para outro fim sem o seu consentimento; d) deve haver uma via para os indivíduos corrigirem ou alterarem registros de informações identificáveis sobre eles e; e) toda organização que cria, mantém, usa ou dissemina registros de informação pessoal deve garantir a confiabilidade e tomar as precauções razoáveis para prevenir mau uso.

Essa síntese de caráter principiológico é decorrente da grande variedade de temas aos quais o problema da privacidade e proteção de dados estava relacionada. Durante os encontros tratou-se de questões relacionadas à criação de bancos de dados, sejam públicos ou privados. Discutiu-se os usos das informações da seguridade social, da saúde e da justiça criminal em outros contextos como agências de emprego ou *bureaus* de crédito.

O artigo de Gellman elenca uma série de críticas às FIPs. Entre elas algumas propostas de que sejam ultrapassados e devam ser substituídos, outras de que estão defasados e precisam ser atualizados. Essas críticas relacionam-se à capacidade de esses princípios constituírem um mecanismo eficiente de defesa de direitos no contexto informacional de *big data* e inteligência artificial. Entre as limitações é comum apontar-se a ausência de uma autoridade para o *enforcement* e de que se baseiam exclusivamente em autorregulação. Entre as atualizações, encontram-se propostas de que incluam revisão humana de decisões automatizadas.

Não cabe neste trabalho discutir a atualidade e adequação desses princípios, pensados em uma época em que os problemas no uso da inteligência artificial preocupavam apenas pequenos grupos de cientistas. É válido destacar, contudo, que os encontros em 1973 já antecipavam problemas que apenas agora recebem a devida atenção. Alguns estudiosos já apontavam, inclusive, a necessidade de se estabelecer

[226] GELLMAN, R. *Fair Information Practices*: A Basic History. v. 2.19. Rochester, NY: Social Science Research Network, 2019. Disponível em: https://papers.ssrn.com/abstract=2415020. Acesso em: 11 jun. 2020.

um devido processo dos dados (*informational due process*) e *accountability*[227]*.* Nesse sentido, nos parece promissora a proposta, já apontada por Kaminski[228], de que comecemos a pensar a questão em termos de *accountability* sistêmica dos algoritmos.

O devido processo legal tem sido apontado como um alicerce importante para o futuro do direito à privacidade e da proteção de dados[229], e é nesse ponto que Crawford e Schutz veem a saída para superar os problemas do uso de dados no contexto de processamento algorítmico em massa em face do mero direito à informação presente na regulação tradicional, como os FIPs.

Como demonstrado por Bennet[230], o conteúdo dos FIPs permite que diversos modelos de regulação da privacidade sejam possíveis, conforme descritos no capítulo cinco de sua obra. Uma das formas de regular a proteção de dados é por meio do controle subjetivo. A garantia aos titulares de um direito ao conhecimento sobre seus dados pode desencorajar seu uso indevido e melhorar a qualidade das informações.

Esse direito à informação implica em ao menos três processos inter-relacionados. O primeiro decorre do direito de saber sobre a existência dos bancos de dados e seus propósitos, que implica, para o controlador de dados, a obrigação de manter pública a informação sobre o banco de dados. O segundo diz respeito aos direitos de um indivíduo saber sobre as informações a ele relacionadas no banco de dados, o que implica para o controlador a obrigação de informar o titular sobre meios de obter as informações. O terceiro — e último — diz respeito ao processo para correção ou eliminação dos dados.

Esse "direito de saber" pode ser encontrado em todas as legislações de proteção de dados. No entanto, encontra limites técnicos, sociais e jurídicos. Como podemos notar, não há nessa concepção clássica de direito à informação a obrigação de vincular transparência à capacidade de intervenção efetiva nos processos de tratamento de dados pessoais presentes nos contextos de *profiling* e *big data*. E é nesse sentido que os FIPs encontram uma limitação. Sem abandonar seus

---

[227] HOOFNAGLE, C. J. *The Origin of Fair Information Practices: Archive of the Meetings of the Secretary's Advisory Committee on Automated Personal Data Systems* (SACAPDS). Berkeley Law, 16 jul. 2014. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2466418. Acesso em: 21 jun. 2021.

[228] KAMINSKI, M. E. The Right to Explanation, Explained. *Berkeley Technology Law Journal*, v. 34, n. 189, 2019. Disponível em: https://papers.ssrn.com/abstract=3196985. Acesso em: 27 maio. 2020.

[229] BIONI, B.; MARTINS, P. Devido processo informacional: um salto teórico-dogmático necessário? *Jota*, 15 jul. 2020. Disponível em: https://www.jota.info/opiniao-e-analise/artigos/devido-processo-informacional-um-salto-teorico-dogmatico-necessario-15072020. Acesso em: 15 jul. 2020.

[230] BENNET, C. Regulating privacy: data protection and public policy in Europe and the United States. Ithaca, New York: Cornell University Press, 1992.

fundamentos, alguns autores defendem que é preciso avançar para uma concepção mais sistêmica, de forma a garantir direitos e possibilitar a reparação de injustiças. A transparência continua sendo uma peça fundamental na regulação da privacidade, contudo, é vista como uma etapa inserida num processo mais amplo.

Mais do que procedimentos, Crawford e Schultz defendem um resgate dos *valores* subjacentes ao devido processo legal, conforme consolidados na doutrina jurídica. Tais valores podem ser identificados como: "(1) precisão; (2) aparência de justiça; (3) igualdade de *inputs* no processo; (4) previsibilidade, transparência e racionalidade; (5) participação; (6) revelação e (7) privacidade-dignidade."[231] Somados a estes, vêm também a noção de juiz imparcial e de separação de poderes, que pode ser aplicada por analogia aos processos de *Big Data* quando se observa que há pouca ou nenhuma regulação das interações entre os atores na cadeia de tratamento, nem há um sistema de freios e contrapesos para evitar que haja vieses no sistema.

A título de síntese, os autores definem o devido processo legal, conforme historicamente desenvolvido no sistema estadunidense, como um requerimento constitucional de que qualquer privação da liberdade ou propriedade de um indivíduo seja precedida, no mínimo, por uma notificação e uma oportunidade de uma audiência diante de um julgador imparcial.

Transportando cada um destes elementos para uma lógica informacional de *Big Data*, os autores fazem as seguintes considerações: no que se refere à notificação, este aviso deve conter o tipo de previsões feitas a partir de determinados dados, as fontes de dados utilizados como *inputs*. No mínimo, defende-se que as pessoas afetadas devem saber quais as questões previstas e, idealmente, os dados utilizados e a metodologia. O aviso também deve proporcionar um mecanismo para acessar a trilha ou registro de auditoria criada no processo preditivo.

No caso da oportunidade, concedida ao titular, de ser ouvido, trata-se de uma segunda etapa que lhe possibilita desafiar determinada decisão tomada com base em *Big Data*, e, se necessário, corrigir informações consideradas imprecisas, e até mesmo se opor a elas caso não concordasse ou lhe impactasse negativamente. Isso incluiria o exame das evidências empregadas para informar uma decisão, como o

[231] REDISH, M.; MARSHALL, L. Adjudicatory Independence and the Values of Procedural Due Process. *Yale Law Journal*, [*S. l.*], v. 95, n. 3, 1986. Disponível em: https://digitalcommons.law.yale.edu/ylj/vol95/iss3/1. Acesso em: 21 jun. 2021.

input de dados e a lógica algorítmica utilizada. A ideia de juiz imparcial, por fim, busca endereçar a falácia de que processos que fazem uso de *Big Data* acabam por gerar resultados neutros e objetivos, bem como, enquanto uma "função de separação de poderes", examinar as relações entre aqueles que desenvolvem *softwares* e sistemas e aqueles que os aplicam para coletar, minerar e gerar dados relevantes.

No mesmo ano, outro trabalho, intitulado "The Scored Society: Due Process for Automated Predictions"[232] abordou a mesma questão, também no marco legal estadunidense, a partir do contexto mais específico de *scoring* automatizado de crédito baseado em mineração de *Big Data*. Nesse segundo *paper*, Citron e Pasquale também clamam por uma construção de devido processo legal aplicado a estas tecnologias e práticas, mas partem de argumentos relacionados à natureza e objetivos da Inteligência Artificial. Segundo os autores, a Inteligência Artificial tem uma perspectiva técnica, advinda da engenharia, que foca em garantir resultados pelo trabalho de máquinas que "[...] requeririam inteligência se operados por seres humanos."[233]

A abordagem cognitiva, por outro lado, foca em desenhar sistemas que operem tal qual uma mente humana. A distinção, embora sutil, é relevante, pois a perspectiva técnica confere maior peso ao resultado, independente da forma com a qual ele é obtido — invoca-se, aqui, a metáfora da "caixa preta", em que *inputs* são convertidos em *outputs* sem que se revele como. Os autores afirmam que sistemas de *scoring* costumam ser abordados por esta perspectiva, como uma tecnologia de gerenciamento de risco, mas que os potenciais prejuízos que implicam para os seres humanos demandam uma retomada da abordagem cognitiva.

Diante da constatação de que os benefícios dos sistemas automatizados de *scoring* são mais dificilmente vislumbrados do que os seus riscos, os autores apostam na construção de um sistema de barreiras, por meio de transparência e mecanismos individuais de contestação e sistêmicos de auditoria, com base em instrumentos e organismos existentes no sistema jurídico-regulatório estadunidense.

Para atingir esta mudança, pensando especificamente na prática de *scoring*, os

---

[232] CITRON, D. K.; PASQUALE, F. The Scored Society: Due Process for Automated Predictions. *Washington Law Review*, v. 89, 2014. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2376209. Acesso em: 2 abr. 2021.

[233] CHOPRA, S.; WHITE, L. F. *A Legal Theory for Autonomous Artificial Agents*. [*S. l.*]: University of Michigan Press, 2011. Disponível em: https://www.jstor.org/stable/10.3998/mpub.356801. Acesso em: 27 maio. 2020.

autores sugerem que, já na fase de coleta de dados, haja constrições semelhantes àquelas estabelecidas pelo *Fair Credit Reporting Act*, como a concessão aos cidadãos de direitos de inspecionar, corrigir e contestar dados imprecisos, além de saber a fonte da qual os dados em questão provêm. Depois, no estágio do cálculo do *score*, idealmente tais cálculos seriam públicos e os processos inspecionáveis. Ainda que os autores reconheçam a validade de se alegar segredo de negócios em casos como esses, clamam por uma mudança de paradigma do sigilo para a transparência. O terceiro estágio diz respeito à comunicação ao titular de dados na hipótese de compartilhamento de dados/análises de dados com terceiros e o quarto, considerado pelos autores mais controverso, refere-se à exigência de processos de auditoria em casos de *profiling* em áreas consideradas sensíveis, como emprego, seguros e saúde.

Diante da inexistência, nos Estados Unidos, de legislação específica que dê suporte a estas propostas, os autores recorrem à figura da F*ederal Trade Commission* (FTC) e seus poderes de supervisão. Nesse sentido, sugerem que a transparência pode ser um caminho para facilitar que a FTC promova auditorias e testes e monitore o mercado para identificar sistemas enviesados e arbitrários. Um segundo passo, para a Comissão, seria a emissão de Relatórios de Avaliação de Impacto para avaliar os impactos negativos de sistemas de *scoring* sobre grupos protegidos, além de caracterizações errôneas, resultados arbitrários e riscos à privacidade, em geral.

A FTC tem desenvolvido uma série de instrumentos sobre as questões de privacidade e proteção de dados nos marcos dos FIPs[234]. Em decorrência da emergência de aplicações de inteligência artificial, o órgão tem se dedicado a avançar na questão da *accountability* algorítmica. Em publicação de abril de 2020, o diretor Andrew Smith faz uma série de recomendações para as empresas que utilizam inteligência artificial[235].

A publicação resgata a sua atuação para decisões automáticas no contexto de análise de crédito e o relatório "Big Data: A Tool for Inclusion or Exclusion?*",* de 2016[236], que trazia os riscos e cuidados para evitar viés e discriminação, apresentando

---

[234] GELLMAN, R. *Fair Information Practices*: A Basic History. v. 2.19. Rochester, NY: Social Science Research Network, 2019. Disponível em: https://papers.ssrn.com/abstract=2415020. Acesso em: 11 jun. 2020.

[235] SMITH, A. Using Artificial Intelligence and Algorithms. *Federal Trade Commission*, 8 abr. 2020. Disponível em: https://www.ftc.gov/news-events/blogs/business-blog/2020/04/using-artificial-intelligence-algorithms. Acesso em: 23 jul. 2020.

[236] FEDERAL TRADE COMMISSION. *Big Data:* A Tool for Inclusion or Exclusion? Understanding the Issues. [*S. l.*]: FTC, 2016. Disponível em: https://www.ftc.gov/reports/big-data-tool-inclusion-or-exclusion-understanding-issues-ftc-report. Acesso em 23 jul. 2020.

uma série de recomendações. Entre elas há, por exemplo, o dever de ser transparente na coleta de dados, sobre finalidades e fontes de informações. Além disso, recomenda que as empresas expliquem as decisões, ao menos sobre quais dados e quais foram os fatores determinantes. Recomenda ainda que se realizem análises para evitar discriminações e que deem uma oportunidade para que os consumidores possam corrigir informações. O documento ainda sugere que as empresas se preocupem com a robustez e a representatividade dos dados usados, bem como que se mantenham padrões de governança e de *accountability*.

No entanto, tais medidas carecem de mecanismos efetivos de *enforcement* e de que sejam estabelecidos critérios claros para o exercício da transparência. Citron e Pasquale defendem que, no mínimo, todo titular de dados deve ter acesso a todas as informações sobre dados relativos a si e, idealmente, a lógica por trás dos sistemas preditivos de *scoring* deveria estar disponível para inspeção pela sociedade. As objeções possivelmente levantadas quanto a este modelo aprofundado de transparência — como empecilhos à inovação, segredos de negócio, propriedade intelectual e a possibilidade de que os objetos das classificações passem a "burlar o sistema" — são contrapostos, pelos autores, aos prejuízos incalculáveis a direitos básicos dos cidadãos sujeitos a estas práticas, de forma que, embora legítimos, devem ser argumentos excepcionais.

Convém apontar como o segredo de negócio tem se mostrado uma barreira no processo de transparência e para a *accountability* dos algoritmos. Em algumas aplicações esse fator pode trazer ameaças sérias aos direitos constitucionais. Na justiça criminal, por exemplo, algoritmos estão relacionados a decisões importantes como cálculo de pena[237] ou mesmo de culpabilidade. A falta de conhecimento sobre o algoritmo torna-se um obstáculo ao direito de defesa, uma vez que não é possível questionar seus resultados.

Um projeto de lei[238], de autoria do deputado californiano Mark Takano, pretende impedir que os juízes neguem à defesa o acesso ao código fonte ou informações sobre

---

[237] O *software* COMPAS, utilizado em vários estados, foi testado por uma agência de jornalismo investigativo. O resultado da análise encontrou fortes evidências de viés racial no cálculo de probabilidade de reincidência. Cf. ANGWIN, J. *et al.* Machine Bias. *ProPublica*, 23 maio 2016. Disponível em: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. Acesso em: 27 maio. 2020.

[238] ESTADOS UNIDOS DA AMÉRICA. *H.R.4368 — 116th Congress (2019-2020)*: Justice in Forensic Algorithms Act of 2019. Washington D.C., 10 fev. 2019. Disponível em: https://www.congress.gov/bill/116th-congress/house-bill/4368/text. Acesso em: 24 jul. 2020.

os *softwares* que produzem evidências contra os réus. O deputado argumenta que qualquer *software* utilizado na justiça criminal deve ser transparente e acessível à defesa, como forma de garantir o devido processo[239].

O projeto ainda prevê a criação de padrões para o desenvolvimento de *softwares* de investigação forense. Os padrões devem incluir uma avaliação de risco, de viés racial e de gênero durante o desenvolvimento do programa. O padrão deve divulgar os princípios científicos e métodos utilizados no processamento, bem como apresentar testes sobre o funcionamento do sistema, incluindo sua precisão, seu grau de confiança, robustez e sua reprodutibilidade. O teste deve incluir ainda a análise da representatividade dos dados utilizados no seu desenvolvimento.

Podemos perceber que além da garantia da transparência vedando o uso de segredos comerciais, o projeto prevê um padrão de governança mais sistêmico. Outras propostas de regulação nos EUA buscam da mesma forma estabelecer um maior grau de controle social sobre o sistema além do controle meramente subjetivo ou da autorregulação[240], em convergência com as propostas de Citron e Pasquale. Esse debate tende a se acentuar em decorrência da decisão em favor do ativista austríaco Max Schrems pela Corte de Justiça da União Europeia[241]. A decisão invalidou o acordo de transferência de dados entre a Europa e os Estados Unidos da América e impõe aos estadunidenses a necessidade de discutir uma maior adequação à GPDR para atuar no mercado europeu e a eventual necessidade de uma lei federal de proteção de dados.

As reflexões teóricas e as propostas de regulação apresentadas têm em comum o pano de fundo de crescente preocupação com as consequências do advento e avanço de sistemas que fazem uso de volumes massivos de dados pessoais e, por sua natureza, implicam riscos substanciais a direitos dos titulares de dados, ao mesmo tempo em que se tornam progressivamente mais difíceis de compreender. Também une os autores a identificação da incapacidade ou insuficiência dos marcos legais vigentes à época da publicação dos *papers*, seja nos Estados Unidos ou na Europa,

---

[239] TAKANO, M. Opening the Black Box of Forensic Algorithms. *Medium*, 3 dez. 2019. Disponível em: https://medium.com/@repmarktakano/opening-the-black-box-of-forensic-algorithms-6194493b9960. Acesso em: 24 jul. 2020.

[240] ESTADOS UNIDOS DA AMÉRICA. *H.R. 2231 — 116th Congress (2019-2020):* Algorithmic Accountability Act of 2019. 4 nov. 2019. Disponível em: https://www.congress.gov/bill/116th-congress/house-bill/2231/all-info. Acesso em: 24 jul. 2020.

[241] MAJOR EU-US data protection agreement struck down. *BBC News*, [*S. l.*], 2020. Disponível em: https://www.bbc.com/news/technology-53418898 . Acesso em: 24 jul. 2020.

e seu foco, dentre outros elementos correlatos, para garantia de *accountability*.

3.1.2 União Europeia

Um conjunto de contribuições, representativo da literatura mais atual sobre o direito à explicação e as diversas questões que o permeiam, revela um certo movimento: verifica-se que textos um pouco mais antigos aos anos que precederam a aprovação da GDPR são mais focados em identificar precisamente as problemáticas decorrentes do avanço de técnicas de decisões automatizadas a partir do uso massivo de dados pessoais, e focam em soluções abrangentes relacionadas à transparência e *accountability,* seja por meio de expectativas relativas à então proposta de novo Regulamento europeu, ou por meio de variações da noção estadunidense de devido processo legal. Com a aprovação da GDPR, estabeleceu-se o debate em torno da existência e alcance de um direito à explicação, conduzido principalmente por Wachter *et al.* e Selbst e Powles[242].

Passado este primeiro momento de intenso debate nos meses que sucederam a aprovação do Regulamento, outras contribuições passaram a questionar um foco excessivo na dualidade da existência ou não existência do direito à explicação e propuseram outras abordagens[243], conforme apresentamos no Capítulo 2, que podem ser divididas em dois grupos: (i) há aquelas que ampliam o olhar sobre a GDPR e seu potencial regulatório e identificam outros elementos que podem mitigar os riscos do cenário atual (como as medidas de "*accountability* sistêmica", que caminham em direção a uma transparência "qualificada" e a um preenchimento do ideal estadunidense do devido processo legal); (ii) e aquelas que focam em questionamentos mais principiológicos, dando "um passo atrás" a partir de ponderações como o que é uma explicação, se explicações de fato são necessárias,

---

[242] WACHTER, S.; MITTELSTADT, B.; RUSSELL, C. Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR. *SSRN Electronic Journal*, v. 31, n. 2, 2017. Disponível em: https://www.ssrn.com/abstract=3063289. Acesso em: 27 maio. 2020; SELBST, A. D.; POWLES, J. Meaningful Information and the Right to Explanation. *International Data Privacy Law*, v. 7, n. 4, p. 233-242, 2017. Disponível em: https://papers.ssrn.com/abstract=3039125. Acesso em: 27 maio 2020.

[243] A despeito disso, é possível afirmar que as tipologias desenhadas pelos autores que estruturam o debate sobre a existência do direito à explicação seguem presentes e influenciando os escritos mais recentes. Em 2019, a fundação Bertelsmann Stiftung disponibilizou um *discussion paper* intitulado: "The General Data Protection Regulation and Automated Decision-making: Will it deliver?" em que faz um sobrevoo sobre todos os dispositivos relevantes da GDPR acerca de decisões automatizadas e os riscos que elas produzem frente aos direitos de titulares de dados pessoais, grupos e a sociedade, como um todo. Dentre as limitações observadas estão algumas daquelas pontuadas por Wachter *et al.*

o valor da transparência, etc.

Um texto que busca impulsionar uma perspectiva pouco adotada na discussão acadêmica sobre o direito à explicação é “The Right to Explanation, Explained”[244], segundo o qual o recorte majoritariamente adotado para este debate obscureceu o significativo regime de *accountability* algorítmica estabelecido pela GDPR. Enquanto nos Estados Unidos haveria um vácuo regulatório, a GDPR, segundo Margot Kaminski, é um exemplo forte (tanto positivo como negativo) de um regime de *accountability* algorítmica vigente. A discussão sobre o direito à explicação, de acordo com ela, pode vir a ser considerada “confusa” para um público estadunidense porque esse debate, da maneira como foi estruturado, obscureceu a real dimensão e profundidade do referido regime.

Para além dos quatro artigos da GDPR (13, 14, 15 e 22) que endereçam diretamente práticas de transparência e processos decisórios automatizados, inclusive *profiling,* previsões como o direito de oposição, o direito de retificação, *data protection by design* e *by default* e a exigência de *Data Protection Impact Assessments* (DPIAs), integram um conjunto de dispositivos que caracterizam este regime, considerado no artigo como consideravelmente mais amplo, forte e profundo do que aquele existente sob a Diretiva anterior.

Debruçando-se sobre o conteúdo do art. 22, referente ao direito do indivíduo não ser submetido a decisões exclusivamente automatizadas, Kaminski destaca a exigência de garantia de salvaguardas ao titular de dados no caso de aplicação de alguma das exceções previstas no dispositivo, dentre as quais explicita-se o direito de “[...] ao menos obter intervenção humana, expressar seu ponto de vista e contestar a decisão.”[245] Para Kaminski, esses requisitos explicitamente criam uma versão de “devido processo legal algorítmico”, por meio do direito a uma oportunidade de ser ouvido. O uso do termo “ao menos” sugere, ainda, que este é apenas um patamar mínimo.

No que se refere aos deveres de notificação e direito de acesso, extraídos dos arts. 13, 14 e 15 da GDPR, relembra que todos estes dispositivos exigem a divulgação da “existência de processos decisórios automatizados, inclusive *profiling*”, e de “informação significativa sobre a lógica envolvida, assim como o significado e as

[244] KAMINSKI, M. E. The Right to Explanation, Explained. *Berkeley Technology Law Journal*, v. 34, n. 189, 2019. Disponível em: https://papers.ssrn.com/abstract=3196985. Acesso em: 27 maio. 2020.
[245] Ibidem, p.

consequências vislumbradas deste tratamento para o titular de dados", e que há diferenças de *timing,* já que os arts. 13 e 14 permitem o acesso à informação quando os dados são obtidos, mas o art. 15 garante esse acesso a qualquer momento. A redação dos artigos e a problemática do *timing* tem levado, conforme explorado nesta revisão, a divergências quanto à natureza das informações exigidas em cada caso: alguns defendem que, no caso dos arts. 13 e 14, trata-se de uma informação mais panorâmica sobre um determinado sistema e que, quanto ao art. 15, poder-se-ia invocar um direito de acesso a informações sobre decisões específicas.

Diante da controvérsia, a autora opta por focar nas interpretações já produzidas pelo Grupo de Trabalho do Art. 29, bem como no conteúdo dos Considerandos da GDPR, que considera o principal norte interpretativo sobre como os parâmetros estabelecidos pela norma devem ser aplicados. Quanto às *guidelines* do Grupo de Trabalho, embora não tenham força de lei, são um forte indicativo de como a lei será efetivamente interpretada por seus aplicadores. Dessa forma, muito embora apenas o texto do Regulamento seja formalmente vinculante, tanto seus Considerandos, quanto às orientações construídas pelos órgãos consultivos devem desempenhar um papel significativo, na prática, em guiar o comportamento das empresas e entidades, o que corresponde à própria concepção da GDPR como um regime de governança colaborativa.

A autora sustenta que, se o texto do Regulamento é repleto de parâmetros abertos, isso decorre de um objetivo maior de que a sua substância seja desenhada ao longo do tempo por meio de um diálogo constante entre reguladores e entidades. Dessa forma, Kaminski sustenta que centrar argumentos sobre a existência e alcance do direito à explicação na natureza (vinculante ou não) dos textos que lhe dão suporte é não apenas insistir em uma tecnicidade pouco relevante, como também negar a natureza evolutiva e colaborativa da GDPR.

Partindo para as interpretações já disponíveis desde a aprovação do Regulamento, o primeiro ponto é que as *guidelines* do Grupo de Trabalho esclarecem que o art. 22 é uma efetiva vedação à tomada de decisões exclusivamente por algoritmos, e não um mero direito de oposição[246], de forma que o seu emprego será

[246] EUROPEAN DATA PROTECTION BOARD. *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 679/2016*. Bruxelas: European Commission, 2016. Disponível em: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053. Acesso em: 02 jul. 2021.

submetido à justificativa de incidência de alguma das três exceções elencadas no artigo. Também esclarece que, para afastar a incidência do art. 22, que fala em processos "exclusivamente" automatizados, o envolvimento humano no sistema deve ser "significativo" e realizado por quem tenha a autoridade e a competência para alterar a decisão em questão, além de ter acesso a informações que vão além dos *outputs* do algoritmo.

No que se refere a outro ponto de recorrente discussão, o significado de "decisão com efeitos legais ou igualmente significativos", a autora aponta que tanto o Considerando 71, quanto às *guidelines,* trazem exemplos concretos, como decisões sobre crédito e *e-recruiting*[247]. Podemos apontar as decisões que afetam circunstâncias financeiras, acesso aos serviços de saúde, emprego, que colocam alguém "em séria desvantagem" ou decisões que afetam o acesso à educação.

As *guidelines* também explicam que alguns tipos de publicidade comportamental podem ser cobertos quando forem particularmente intrusivos, contra titulares particularmente vulneráveis. Além disso, práticas de precificação também podem disparar a aplicação do art. 22 se "[...] preços proibitivos efetivamente impedem que pessoas tenham acesso a bens e serviços."[248] Quanto à questão de segredos de negócio, as *guidelines* esclarecem que, ainda que se trate de uma preocupação válida, empresas não podem se apoiar nessa proteção para negar acesso ou se recusar a fornecer informações. As *guidelines* também explicam que as exceções ao artigo (contratual e por meio do consentimento) têm interpretação restritiva — deve haver uma demonstração de adequação e necessidade que justifique o emprego de processos automatizados, mesmo nesses casos. Quanto ao consentimento, especificamente, as *guidelines* determinam que os indivíduos devem receber informação suficiente sobre as consequências do *profiling*, de forma que seu consentimento possa ser considerado uma escolha informada.

Finalmente, quanto ao art. 22 e o debate sobre as salvaguardas adequadas

---

[247] Recital 71: "[…] such as automatic refusal of an online credit application or e-recruiting practices without any human intervention…in particular to analyse or predict aspects concerning the data subject's performance at work, economic situation, health, personal preferences or interests, reliability or behaviour, location or movements, where it produces legal effects concerning him or her or similarly significantly affects him or her." (UNIÃO EUROPEIA. *Regulamento Geral sobre a Proteção de Dados — Considerando 71*. Disponível em: https://gdpr-text.com/pt/read/recital-71/. Acesso em: 21 jun. 2021).

[248] EUROPEAN DATA PROTECTION BOARD. *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 679/2016*. Bruxelas: European Commission, 2016. Disponível em: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053. Acesso em: 02 jul. 2021.

para legitimar decisões baseadas em processos automatizados, a autora destaca que, além da menção expressa a um direito à explicação no Considerando 71, que trata do tema, as *guidelines* empregam o mesmo vocabulário em ao menos três ocasiões. O raciocínio por trás desta inclusão é que, para que o indivíduo possa desafiar uma decisão em particular e expressar seu ponto de vista, ele precisa primeiro entender como essa decisão foi atingida e com base em quê.[249] Em outras palavras, um indivíduo tem direito à explicação porque ele é necessário para o exercício de outros direitos enumerados expressamente na GDPR, tais como o direito de contestação e expressão de ponto de vista sobre uma decisão.

A partir destas breves considerações, Kaminski sustenta que as referidas *guidelines* estabelecem, para além de um direito à explicação, uma versão de devido processo legal algorítmico. Para além do aspecto individual do devido processo (o direito a uma oportunidade de ser ouvido), elas interpretam elementos do Regulamento, como a exigência de "salvaguardas adequadas", como também incluindo medidas de *accountability* sistêmica, como auditorias e formação de Conselhos de Ética. Estas medidas podem ser interpretadas tanto como uma forma de *oversight* e *enforcement* para garantia de direitos individuais ou como parte de um regime de governança colaborativa. Retornando ao seu argumento mais abrangente, a autora pontua que, se os Considerandos e *guidelines* não eliminaram por completo disputas interpretativas em torno do direito à explicação, eles certamente esclarecem as previsões da GDPR sobre transparência e *accountability* algorítmica e tornam-as mais rigorosas.

Nesse sentido, o art. 22 é mais amplo do que o seu correspondente na Diretiva — ele não se limita ao *profiling* individual, incluindo outros tipos de decisões automatizadas, pois o artigo tem uma interpretação mais protetiva do nível de envolvimento humano no processo e não é considerado um direito de mera oposição. A GDPR, por si só, também é uma regulação mais forte e potente do que a Diretiva, na medida em que confere papel de destaque a *Data Protection Officers* (DPOs), *Data Protection Impact Assessments* (DPIAs) e aos princípios de *data protection by design.* Kaminski defende que a GDPR é um passo significativo em direção ao que Pasquale denomina "transparência qualificada", isto é, "[...] um sistema de revelações de diferentes graus de profundidade e abrangência, destinadas a diferentes

[249] UNIÃO EUROPEIA. *Regulamento Geral sobre a Proteção de Dados — Considerando 71*. Disponível em: https://gdpr-text.com/pt/read/recital-71/. Acesso em: 21 jun. 2021.

receptores”[250], uma combinação de medidas individuais e sistêmicas.

Tal entendimento contrapõe-se à ideia, muito ventilada, de uma “falácia da transparência”, que a autora rejeita por entender que se trata de uma representação errônea da essência desta transparência. De acordo com ela, no debate sobre transparência, o receptor de uma informação definirá o seu conteúdo, *timing* e modo de apresentação. Assim, provisões individuais têm como objetivo empoderar indivíduos para que invoquem seus direitos sob a GDPR, caso em que não faria sentido, por exemplo, fornecer tal informação na forma de códigos-fonte ou fórmulas matemáticas, que pouco teriam a dizer a este indivíduo.

A ideia de “inteligibilidade” da explicação concedida ao indivíduo foi muito confundida por determinados autores como uma forma de reduzir o nível de informação, por simplificá-la, mas Kaminski defende que essa postura tem como objetivo evitar que o “excesso de transparência”, isto é, o volume excessivo e confuso das informações, acabe representando um fator de obscurecimento e dificuldade para o exercício dos direitos de contestação, correção e apagamento. Essa lógica, para a autora, demonstra uma característica interessante da transparência: que a sua substância é determinada pela substância de outros direitos que lhe são subjacentes. Em outras palavras, se o indivíduo tem, por exemplo, um direito de correção, é preciso que primeiro ele veja o erro.

Isso não significa, segundo a autora, que o direito individual à explicação e os direitos de transparência, conforme garantidos pela GDPR, assegurem um direito a toda e qualquer informação sobre um algoritmo. Por outro lado, as medidas de *accountability* sistêmica significam que, se o indivíduo não tem (e não precisa) ter acesso a códigos-fonte e outras informações complexas, outros atores do sistema podem ter e há veículos apropriados para que estas informações sejam apresentadas de uma forma dialógica, seja por meio de *Data Protection Impact Assessments* (DPIAs) ou, de forma mais genérica, pelos poderes de execução e supervisão das autoridades nacionais.

A dinâmica entre medidas individuais de promoção de transparência e medidas sistêmicas de *accountability* tem, de acordo com Kaminski, um aspecto temporal: as primeiras são mais limitadas, exercidas em momentos específicos (coleta,

[250] PASQUALE, F. *The Black box society:* the secret algorithms that control money and information. First Harvard University Press paperback edition ed. Cambridge, Massachusetts; London, England: Harvard University Press, 2015.

consentimento, decisão) e, por isso, raramente atingirão os momentos anteriores à concepção e o completo desenvolvimento de um algoritmo. A lógica sistêmica, por outro lado, tem como objetivo ser contínua e dinâmica, portanto, mais duradoura.

3.1.2.1 GDPR

Pudemos notar que muitos autores defendem a existência do direito à explicação na GDPR baseando-se nos arts. 13, 14 e 15 do Regulamento. Os dispositivos versam sobre as informações que devem ser fornecidas ao titular. Todavia, essa interpretação tem sido debatida por especialistas, que chegam a conclusões distintas. Os debates opuseram aqueles que, por um lado, argumentam que a ausência do termo “explicação” no texto da GDPR não permitiria afirmar categoricamente a existência de um direito à explicação. De outro lado, estão aqueles que argumentam que tal direito poderia ser inferido a partir desses artigos, ao se referirem ao dever de fornecer informações significativas sobre a lógica envolvida na tomada de decisões automatizadas, quando estas impactarem a vida dos titulares dos dados.

No contexto de decisões automatizadas, o direito à explicação encontra respaldo no art. 13 da GDPR, que determina que devem ser fornecidas ao titular dos dados informações significativas sobre a lógica do processamento automatizado, bem como sobre o significado e as consequências previstas para o titular dos dados do processamento. A explicação sobre a lógica envolvida no tratamento dos dados pessoais é, de certa forma, um esclarecimento sobre o que será feito com tais dados, ou seja, um direito à explicação. Neste contexto, a consideranda 71 da GDPR afirma que tais direitos incluiriam: (i) o de obter uma explicação referente à decisão tomada; e (ii) o de desafiar essa decisão.

Em qualquer dos casos, tal tratamento deverá ser acompanhado das garantias adequadas, previstas no art. 22(3), que deverão incluir a informação específica ao titular dos dados e o direito de obter a intervenção humana, de manifestar o seu ponto de vista, de obter uma explicação sobre a decisão tomada na sequência dessa avaliação e de contestar a decisão.

Assim, apesar do termo “explicação” não estar previsto no corpo do

Regulamento, apenas nas suas considerandas, que não são vinculantes[251], é possível argumentar que há um direito à explicação no Regulamento Europeu de Proteção de Dados. Ele teria como base o princípio da transparência e o direito de acesso aos dados, que incluiria o direito a receber explicação sobre a lógica subjacente de decisões totalmente automatizadas com impacto na vida dos indivíduos – principalmente as que incluem perfis comportamentais.

Do outro lado do Atlântico, a discussão sobre práticas preditivas sobre dados pessoais, avanço das tecnologias que lhes dão sustentação e demandas regulatórias continuou a se estruturar em torno do novo Regulamento Geral de Proteção de Dados europeu, cuja redação direcionou o debate para o direito à explicação, objeto desta tese. Por isso, convém retomarmos em maior profundidade a discussão em torno da legislação europeia.

Cerca de dois meses depois da aprovação do Regulamento, o *paper* "European Union regulations on algorithmic decision-making and a "right to explanation"[252] deu início à discussão que vem ganhando novos contornos e se estende até o momento, dois anos após a entrada em vigor do Regulamento. O argumento de Goodman e Flaxman é simples: o art. 22 do Regulamento, que diz respeito a decisões automatizadas, inclusive *profiling*, pode, a depender de sua interpretação e *enforcement*, exigir uma completa revisão de técnicas algorítmicas amplamente utilizadas e padronizadas na indústria. Mais do que isso, a escolha do Regulamento em conferir o direito aos cidadãos de "receber uma explicação para decisões por algoritmos evidencia a crescente relevância da possibilidade de interpretação humana no *design* de algoritmos."[253] A consequência, naturalmente, é uma crescente necessidade de algoritmos que possam operar efetivamente dentro deste novo marco legal.

Em resumo, o art. 22, parágrafo 1, veda qualquer, com algumas exceções, "decisão baseada exclusivamente em tratamento automatizado, inclusive *profiling*", desde que ela "afete significativamente" (sendo este efeito jurídico ou não) o titular. O

---

[251] VEALE, M. *Governing Machine Learning that Matters*. 2019. Doctoral Thesis (PhD in Science, Technology, Engineering and Public Policy) — University College London, London, 2019. Disponível em: https://discovery.ucl.ac.uk/id/eprint/10078626/1/thesis_final_corrected_mveale.pdf. Acesso em 15 dez. 2020. p. 111-112.

[252] GOODMAN, B.; FLAXMAN, S. European Union regulations on algorithmic decision-making and a "right to explanation". *AI Magazine*, [*S. l.*], v. 38, n. 3, p. 50–57, 2017. Disponível em: https://arxiv.org/abs/1606.08813. Acesso em: 15 dez. 2020.

[253] Ibidem, p. ....

parágrafo 2 elenca as seguintes exceções: quando este tratamento for necessário à celebração ou execução de um contrato, quando houver previsão expressa por legislação dos Estados-membros ou com base em consentimento explícito. Em seguida, o parágrafo 3 estipula que, mesmo quando forem aplicáveis as referidas exceções, os controladores deverão assegurar "salvaguardas apropriadas", inclusive "o direito de obter intervenção humana… expressar seu ponto de vista e contestar a decisão".[254]

Abordando especificamente o direito à explicação em seção própria, Goodman e Flaxman discorrem também sobre os arts. 13 a 15 do Regulamento, que estabelecem deveres de notificação aos controladores e o direito de acesso ao titular e, segundo os autores, complementam as previsões do art. 22, que não se aprofunda sobre o significado de "salvaguardas apropriadas" para além de "intervenção humana". Os arts. 13 e 14 afirmam que, em caso de *profiling,* o titular de dados pessoais tem direito a receber "informação significativa sobre a lógica envolvida", previsão que deu origem a um amplo debate sobre seu significado e alcance.

Um dos termos do debate diz respeito à variedade de algoritmos de inteligência artificial. Em alguns casos de *machine learning,* tal lógica não é conhecida de antemão. O que se opera é justamente a descoberta de lógicas subjacentes aos dados e anteriormente desconhecidas ao programador. Dessa maneira, pode não ser possível explicar em termos compreensíveis qual a lógica envolvida.

Costuma-se dividir os modelos de *machine learning* entre supervisionados e não supervisionados. Os primeiros trabalham com variáveis as quais foram atribuídos valores (*labels*) e obtém um resultado a partir da análise de extensas quantidades desses dados. Os modelos não supervisionados, por sua vez, trabalham com dados sem valoração e procuram justamente por categorias ou lógicas presentes nos dados, desconhecidas para os programadores.

Esses algoritmos de *machine learning* não operam mediante uma lógica de causalidade e explicação tradicional, mas de mera correlação. Diante disso, constatações sobre a "caixa preta" algorítmica levam, conforme reconhecem Goodman e Flaxman, a exigências por mais transparência, sem que necessariamente fique claro o que exatamente se está exigindo.

---

[254] GOODMAN, B.; FLAXMAN, S. European Union regulations on algorithmic decision-making and a "right to explanation". *AI Magazine*, [*S. l.*], v. 38, n. 3, p. 50–57, 2017. Disponível em: https://arxiv.org/abs/1606.08813. Acesso em: 15 dez. 2020.

O texto faz uma importante referência aos tipos de obstáculos à transparência, ou opacidade, enumerados por Burrell[255]: i) opacidade decorrente de estratégia corporativa; ii) opacidade decorrente da especialização da escrita e leitura dos códigos; e iii) opacidade decorrente da incapacidade de compreensão humana da análise multidimensional dos algoritmos, conforme abordamos no Capítulo 2.

Os autores alertam que, embora as discussões sobre explicação se destinem ao primeiro e ao segundo obstáculo, a questão relativa à discrepância entre a linguagem deste tipo de sistema de algoritmos e os limites da cognição humana ainda representa um desafio significativo a ser trabalhado pela indústria.

O artigo não chega a avançar neste ponto, apenas sugere que, se um algoritmo só pode ser explicado quando o seu modelo puder ser articulado e compreendido por um ser humano, logo é razoável supor que qualquer explicação deve, no mínimo, fornecer informações sobre em que medida *inputs* se relacionam com previsões feitas pelo sistema, permitindo que se responda perguntas como as seguintes: "É mais ou menos provável que o modelo recomende um empréstimo se o requerente for parte de uma minoria social? Quais recursos desempenham um papel mais relevante nesta previsão?" Em suma, esta contribuição de Goodman e Flaxman estabeleceu alguns dos termos do debate sobre o direito à explicação, que foram desenvolvidos com profundidade em outros trabalhos.

O primeiro deles, e o mais consistente desafio à existência legal e viabilidade técnica do direito à explicação, é o emblemático artigo "Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation".[256] A posição de Wachter, Mittelstadt e Floridi é que, a despeito do furor em torno da noção de um direito à explicação a partir da aprovação do Regulamento (e particularmente de um direito à explicação sobre decisões específicas), os arts. 13 a 15 do Regulamento só garantiriam informações "significativas, mas limitadas" sobre a lógica de tratamento de dados pessoais nestes contextos e das consequências vislumbradas para o titular. Além disso, consideram ambíguo e limitado o escopo do art. 22, o que indicaria uma falta de precisão linguística e de propósito para a almejada

---

[255] BURRELL, J. *How the Machine "Thinks:" Understanding Opacity in Machine Learning Algorithms*. Rochester, NY: Social Science Research Network, 2015. Disponível em: https://papers.ssrn.com/abstract=2660674. Acesso em: 27 maio. 2020.

[256] WACHTER, S.; MITTELSTADT, B.; FLORIDI, L. *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation*. Rochester, NY: Social Science Research Network, 2016. Disponível em: https://papers.ssrn.com/abstract=2903469. Acesso em: 27 maio. 2020.

proteção dos titulares de dados pessoais contra processos decisórios automatizados.

O primeiro ponto examinado por Wachter *et al.* é o conjunto de significados possíveis para "explicação" quando se fala em direito à explicação. A divisão proposta é entre explicações sobre funcionalidades de um sistema, que incluem a "lógica", "significado" e "consequências vislumbradas", e explicações sobre decisões específicas, que incluiriam as razões e circunstâncias individuais de uma decisão automatizada. Além disso, as explicações também podem ser *ex ante*, isto é, anteriores a uma decisão e, nesse sentido, relacionadas às funcionalidades de um sistema, ou *ex post,* potencialmente relativas tanto à funcionalidade como a decisões específicas.

A partir disso, os autores sustentam que não há direito à explicação na GDPR. Isso porque, da análise das três bases que dariam forma a este direito, extrai-se que sua maior força deriva não do próprio texto do Regulamento, mas sim dos Considerandos que o acompanham. As referidas bases seriam as seguintes: as salvaguardas para o emprego de decisões automatizadas, presentes no art. 22(3) e no Considerando 71[257]; os deveres de notificação estabelecidos pelos arts. 13 e 14 e Considerandos 60-62; e o direito de acesso, garantido pelo art. 15 e Considerando 63.

Quanto às salvaguardas, o argumento dos autores é simples: o Regulamento assegura o direito de obter intervenção humana, expressar pontos de vista e contestar uma decisão, mas não de obter qualquer tipo de explicação sobre ela. O suposto direito à explicação seria mencionado expressamente apenas no Considerando 71, que, *caso fosse vinculante* (o que, como ressaltam os autores, não é), geraria um direito à explicação *ex post* sobre decisões específicas. Acerca desta omissão da menção explícita a um direito à explicação, o *paper* sugere ter sido intencional, já que o texto do Considerando é praticamente idêntico ao do art. 22, à exceção deste único trecho sobre o direito à explicação.

Em relação aos deveres de notificação, a leitura conjunta dos arts. 13 e 14 do Regulamento cria a obrigação de informar ao titular de dados acerca da existência de decisões automatizadas a seu respeito e de "informação significativa" sobre a lógica envolvida e as consequências vislumbradas por este tratamento. Os autores

[257] "[...] In any case, such processing should be subject to suitable safeguards, which should include specific information to the data subject and the right to obtain human intervention, to express his or her point of view, to obtain an explanation of the decision reached after such assessment and to challenge the decision." (UNIÃO EUROPEIA. *Regulamento Geral sobre a Proteção de Dados — Considerando 71*. Disponível em: https://gdpr-text.com/pt/read/recital-71/. Acesso em: 21 jun. 2021).

argumentam que, a despeito de outras sugestões, estes dispositivos não justificam o argumento por um direito à explicação referente a decisões específicas, na medida em que a notificação logicamente *precede* o processo decisório. Além disso, alertam que o *link* entre o art. 22(3), que se refere às salvaguardas, e estes artigos não pode ser extraído literalmente do texto do Regulamento, mas apenas de uma leitura de natureza teleológica ou sistemática.

Essa interpretação, de acordo com Wachter *et al*., é corroborada pelo próprio vocabulário dos artigos, que sugere uma informação genérica sobre como os sistemas funcionam (“lógica envolvida”) e uma previsão de impacto (“consequências vislumbradas”), *antes* de qualquer decisão.

Quanto à terceira possibilidade, de derivação de um direito à explicação do direito de acesso contido no art. 15, os autores analisam as consequências da mudança sutil de um dever de notificação para um direito de acesso sobre o *timing* das explicações. Embora a redação deste artigo seja praticamente idêntica a dos anteriores, o direito de acesso tem a especificidade de depender de requisição do titular e, por essa razão, não tem prazo, o que tornaria plausível a ideia de um direito à explicação *ex post.* Ainda assim, o *paper* sustenta que é razoável duvidar que tal explicação seria exigível em relação a decisões específicas, uma vez que o artigo, a exemplo dos que o precedem, mantém o vocabulário sobre “consequências vislumbradas”, o que sugere um momento futuro, em relação à concessão da informação ao titular.

Diante dessas incertezas provenientes do texto do Regulamento, os autores sustentam que o art. 15 apenas garante um direito de acesso a informações e explicações sobre a funcionalidade de um sistema. A fim de corroborar com esta tese, os autores fazem um resgate histórico da implementação da Diretiva anterior à GDPR, cujo dispositivo correspondente foi aplicado, ao longo dos anos, de forma limitada a este tipo de explicação genérica, por conta, majoritariamente, de questões relacionadas a segredos de negócio.

Feitas estas considerações, o artigo afirma que, embora não haja um direito à explicação na GDPR, a contribuição de um direito desta espécie para a transparência e *accountability* de processos decisórios automatizados poderia justificar um esforço pela sua construção futura. Entretanto, ainda que esse fosse o caso e existisse de fato um direito à explicação relacionado a decisões específicas, outras previsões da própria GDPR restringiriam o seu escopo significativamente. A primeira limitação,

segundo o texto, é a definição restritiva de processo decisório automatizado no art. 22,1, que requer “efeitos legais ou outros efeitos significativos” de decisões baseadas “exclusivamente” em processos automatizados, ambos termos relativamente vagos e que requerem interpretações futuras.

Uma outra limitação encontrada pelos autores, com base em análises da discussão sobre o direito de acesso no âmbito da Diretiva, é a questão dos interesses do controlador de dados, o que é corroborado pelo Considerando 63, segundo o qual o direito de acesso não deve infringir os direitos e liberdades de terceiros, inclusive controladores. Além destas restrições, o *paper* aponta outras exclusivas a um direito supostamente derivado do art. 22(3). Primeiro, remete às exceções à vedação de decisões exclusivamente automatizadas para afirmar que a exigência de salvaguarda definida se aplica apenas no caso de contratos e consentimento explícito, mas não no caso de exceções baseadas em leis aprovadas por membros da União Europeia. Além disso, no caso da exceção baseada em contratos, não haveria uma definição clara de quando este tipo de decisão automatizada seria “necessária” à celebração ou cumprimento de um contrato, o que, de acordo com Wachter *et al.*, pode relegar esta definição exclusivamente ao controlador.

Adiante, o *paper* levanta um tema de extrema relevância para o debate sobre o direito à explicação e especificamente acerca do art. 22, que é a possibilidade de interpretá-lo como uma vedação ou como um direito de oposição, ambiguidade já presente à época da Diretiva e que levou a implementações em ambos os sentidos por diferentes Estados-Membros. As consequências de uma ou outra interpretação são relevantes: se uma proibição, controladores não seriam capazes de tomar decisões com base em processos exclusivamente automatizados sobre um titular a não ser que preenchessem uma das três referidas exceções e garantissem as salvaguardas correspondentes; se um direito de oposição, a restrição ocorreria apenas mediante uma objeção ativa por parte do titular.

Diante de todas as constrições e ambiguidades sugeridas pelos autores, eles optam por invocar um “direito a ser informado” sobre a existência de decisões automatizadas e funcionalidades dos sistemas empregados. A título de conclusão, fazem uma série de recomendações que reputam necessárias para a construção e consolidação de um direito à explicação mais substantivo: (i) a inclusão expressa de um direito de explicação no art. 22(3), que é vinculante; (ii) esclarecimento sobre os termos “existência de”, “informação significativa”, “lógica envolvida”, “significado” e

“consequências vislumbradas” no art. 15,1, h; (iii) esclarecimento da linguagem do art. 22(1) para indicar quando decisões são consideradas baseadas “exclusivamente” em processos automatizados; (iv) esclarecimento da linguagem do art. 22(1) para indicar o que contaria como “efeito legal ou efeito igualmente significativo”; (v) esclarecimento da linguagem do art. 22(2)(a), em relação à “necessidade” de celebração ou cumprimento de contrato; (vi) esclarecimento da linguagem do art. 22 para indicar que se trata de uma vedação; (vii) como contrapartida a segredos de negócio, introdução de mecanismos de auditoria externa ou interna; (viii) apoio a pesquisas futuras sobre a viabilidade de mecanismos alternativos de *accountability*.

Em complemento a este ensaio seminal, Wachter, Mittelstadt e Russell produziram mais dois *papers.* O primeiro, “Counterfactual Explanations Without Opening the Black Blox”[258], tem como premissa as conclusões defendidas no artigo anterior e vai além, adentrando na esfera técnica para adicionar outras limitações à efetivação de um direito à explicação. As três novas assertivas defendidas nessa contribuição são: o desafio de ordem técnica para explicações da racionalidade de um algoritmo em casos concretos; a possibilidade de que o produto destas explicações seja de pouco valor efetivo para o titular de dados; por fim, o artigo destaca que controladores têm um interesse em não compartilhar detalhes dos seus algoritmos para evitar revelar segredos de negócio, violar direitos e interesses de terceiros e permitir que titulares de dados pessoais manipulem os seus sistemas.

A despeito de todas estas limitações, os autores sustentam que o valor da explicabilidade de sistemas de decisões automatizados persiste, mas que um ponto que consideram central tem sido negligenciado neste debate: que uma explicação de decisões automatizadas, e um direito geral de explicação, não dependem, necessariamente, de que o público *entenda* como um algoritmo funciona.

Em outras palavras, o objetivo do artigo é demonstrar que, *a priori*, explicações podem ser fornecidas sem que se abra a “caixa preta”, a partir da proposta de três objetivos principais que uma explicação deve cumprir: (i) informar e auxiliar o titular a compreender por que uma determinada decisão foi atingida; (ii) fornecer a base para a contestação de uma decisão; e (iii) compreender o que pode/deve ser alterado para que um resultado diferente seja obtido no futuro. A solução apresentada pelos autores

[258] WACHTER, S.; MITTELSTADT, B.; RUSSELL, C. Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR. *SSRN Electronic Journal*, v. 31, n. 2, 2017. Disponível em: https://www.ssrn.com/abstract=3063289. Acesso em: 27 maio. 2020.

para que todos os três objetivos sejam cumpridos consiste em fornecer “explicações contrafactuais”, isto é, um raciocínio construído a partir de orações condicionais em que uma delas é falsa.

Diferentemente da lógica que permeia a literatura sobre a explicação no contexto de sistemas automatizados, essa proposta baseia-se em elementos externos que conduzem a uma decisão e não sua lógica interna (“se fator X fosse diferente, então determinada classificação de um indivíduo seria Y”). Conforme explorado minuciosamente no primeiro artigo, os autores defendem que a GDPR não cria um direito que requeira o destrinchamento dos sistemas de decisões automatizadas, de forma que a proposta de abordagem deste *paper* se amoldaria às exigências dos seus dispositivos.

O art. 12 (7)[259] do Regulamento, segundo os autores, corrobora esta tese, na medida em que esclarece que o objetivo dos arts. 13 e 14 é “de forma visível, inteligível e legível oferecer um panorama significativo do tratamento pretendido”. O art. 12(1), por sua vez, estabelece que toda comunicação e informação destinada ao titular de dados deve ser fornecida de maneira “concisa, transparente, inteligível e facilmente acessível”, o que sugere que uma abordagem complexa e baseada em “juridiquês” ou explicações altamente técnicas seria considerada inadequada.

Um terceiro passo na trajetória de Sandra Wachter e pesquisadores parceiros em relação ao direito à explicação no âmbito do Regulamento se deu na forma do artigo “*A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI*”, que reafirma a inexistência de um direito à explicação, assim como seu escopo limitado, caso existisse de forma vinculante na GDPR[260]. A partir disso, o artigo envereda por outro caminho, na medida em que considera que o fornecimento de explicações, quaisquer que sejam suas naturezas, é apenas um dos caminhos

[259] “Art. 12. Transparent information, communication and modalities for the exercise of the rights of the data subject. […] 7. The information to be provided to data subjects pursuant to Articles 13 and 14 may be provided in combination with standardised icons in order to give in an easily visible, intelligible and clearly legible manner a meaningful overview of the intended processing. Where the icons are presented electronically, they shall be machine-readable.” (UNIÃO EUROPEIA. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance)*. Official Journal of the European Union, 4 maio 2016. Disponível em; https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng. Acesso em: 27 ago. 2021).

[260] WACHTER, S.; MITTELSTADT, B. A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI. *LawArXiv*, 12 out. 2018. Disponível em: https://doi.org/10.31228/osf.io/mu2kf. Acesso em: 27 maio 2020.

possíveis para garantir a *accountability* de processos decisórios com base em algoritmos. Isso, porque, enquanto uma explicação é capaz de informar indivíduos sobre resultados de uma decisão e as suposições, previsões e inferências subjacentes que levaram a esta decisão, elas não podem garantir que tal decisão, ou mesmo as suposições, previsões e inferências sejam *justificadas*.

O artigo, que também foi abordado em um recente *talk* no Oxford Internet Institute ("OII")[261] , tem como cerne a problemática da inferência e discorre sobre as diferenças entre inferências feitas por seres humanos e inferências feitas com auxílio de *Big Data* e técnicas de Inteligência Artificial: no último caso, há um volume muito maior de dados sendo processados e não são apenas informações concedidas ativamente, mas todo o "rastro" de dados deixado pelo titular; as decisões extraídas destes dados são contraintuitivas e muitas vezes inexplicáveis do ponto de vista do titular; estes dados têm um ciclo de vida muito mais extenso, podem não ser deletados ou "esquecidos". Os autores sustentam que as referidas diferenças já foram endereçadas por meio do direito ao esquecimento[262], a partir do qual traçam um paralelo para defender que os indivíduos possam ter um direito sobre "como somos vistos".

De acordo com Wachter *et al*., a particularidade das inferências e previsões é que elas não são verificáveis concretamente, o que levanta questões jurídicas relevantes. A primeira delas é se inferências podem ser consideradas dados pessoais, caso em que as proteções criadas pela GDPR poderiam ser invocadas. Nesse caso, os direitos de acesso e de retificação seriam instrumentos relevantes para proteger o titular de dados contra determinados efeitos prejudiciais da referida lógica de inferências subjacente. A Opinião nº 4 do Grupo de Trabalho do Art. 29[263] apresenta um conceito bastante alargado de dado pessoal, que incluiria inferências.

A Corte Europeia de Justiça, por outro lado, tem proferido decisões em outro

---

[261] OII LONDON LECTURE. Show me your data and I'll tell you who you are. 30 out. 2018. Disponível em: https://www.oii.ox.ac.uk/videos/oii-london-lecture-show-me-your-data-and-ill-tell-you-who-you-are/. Acesso em: 21 jun. 2021.

[262] MAYER-SCHÖNBERGER, V. *Delete*. [*S. l.*]: Princeton University Press, 2009. Disponível em: www.jstor.org/stable/j.ctt7t09g . Acesso em: 27 jul. 2020.

[263] ARTICLE 29 DATA PROTECTION WORKING PARTY. *Opinion 04/2014 on surveillance of electronic communications for intelligence and national security purposes.* Bruxelas: European Commission, 10 abr. 2014. Disponível em: https://ec.europa.eu/justice/article-29/documentation/opinionrecommendation/files/2014/wp215_en.pdf. Acesso em: 27 jul. 2020.

sentido. A primeira decisão[264] apresentada por Wachter *et al.* afirma que análises não podem ser consideradas dados pessoais, pois esta seria uma qualificação restrita a fatos e, ainda, que o objetivo da proteção de dados pessoais não é garantir transparência. Outra decisão[265], mais recente, afirma que o conceito de dado pessoal deve ter uma interpretação abrangente, conforme desenvolvido na referida Opinião nº 4. Entretanto, nesta mesma decisão, também se entendeu que a definição sobre a aplicação dos direitos relativos à proteção de dados pessoais, o contexto deve ser observado, partindo-se de uma abordagem teleológica.

Assim, por exemplo, no caso de respostas de um exame, apesar de serem consideradas dados pessoais (pois são usadas para avaliar alguém), evidentemente não seria possível exercer o correspondente direito de retificação. No caso de comentários de avaliadores, por outro lado, entendeu-se que são dados pessoais da pessoa avaliada, mas também do avaliador, de forma que o direito de acesso poderia sofrer determinadas restrições. No caso de exercício do direito de retificação, ainda, seria possível verificar se todo o conteúdo fora corrigido, se todas as folhas estavam corretas, se a prova não fora confundida com a de outro aluno, mas nunca avaliar se os comentários e a avaliação do corretor em si estão corretos.

Com estes exemplos, os autores pretendem demonstrar uma certa incerteza, do ponto de vista de diferentes órgãos com relevantes funções interpretativas, em relação a questões básicas sobre inferências e dados pessoais.

A próxima questão levantada é se existe uma proteção contra inferências feitas a partir de informações sensíveis. Na GDPR, a definição de dados sensíveis claramente abrange inferências, posição também adotada pelo Grupo de Trabalho do Art. 29. A Corte Europeia, por outro lado, em julgamento de 2012, entendeu em outro sentido, propondo uma interpretação restritiva de dado sensível que exige intencionalidade e confiabilidade, ambos elementos que Wachter *et al.* apontam como irrelevantes em uma era de *Big Data*.

A terceira questão levantada no artigo diz respeito às inferências e aos segredos comerciais. A GDPR tem dois objetivos claros: a proteção de dados

---

[264] UNIÃO EUROPEIA. Tribunal de Justiça da União Europeia (3. Câmara). *Casos C-141/12 & 372/12, YS, M e S v. Minister voor Immigratie, Integratie en Asiel*, Data de Julgamento: 17/06/2014. Disponível em: https://curia.europa.eu/juris/liste.jsf?num=C-141/12&language=en. Acesso em: 21 jun. 2021.

[265] UNIÃO EUROPEIA. Tribunal de Justiça da União Europeia (2. Câmara). *Caso C-434/16, Peter Nowak v. Data Prot. Comm'r,* Data de Julgamento: 20/12/2017. Disponível em: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A62016CJ0434. Acesso em: 21 jun. 2021.

pessoais e a remoção de obstáculos ao livre fluxo de dados na Europa. A própria redação do artigo referente ao direito de acesso deixa claro que ele só pode ser exercido quando não comprometer os direitos e liberdades de terceiros, como segredo de negócio e propriedade intelectual.

Diante dessas considerações, que pretendem ir além da discussão sobre um direito à explicação e todas as limitações a ele associadas por esta parcela da literatura sobre o tema, os autores propõem, como diretrizes gerais, uma remodelagem da lógica das legislações de proteção de dados pessoais, com foco em resultados e não em dados de *input* e com a mensagem subjacente de que se trata de uma proteção a pessoas e a sua privacidade, e não apenas a dados pessoais.

Além disso, apresentam uma proposta inicial de "direito a inferências razoáveis", com a ressalva de não se tratar de uma solução *"*one size fits all*"*, mas como uma combinação entre um direito *ex ante* de justificação sobre métodos e bases de dados relevantes, precisos e confiáveis para determinada decisão e um direito de contestação *ex post* que permita o desafio a decisões e inferências eventualmente consideradas desarrazoadas.

Após a contribuição de Wachter *et al.* sobre a suposta inexistência do direito à explicação no âmbito da GDPR, a primeira resposta de fôlego em defesa deste direito veio em 2017 com o artigo "Meaningful Information and the Right to Explanation"[266], de Powles e Selbst, centrado na ideia de que, se é verdade que não existe uma única previsão expressamente rotulada "direito à explicação" na GDPR, ao mesmo tempo também não se trata de um "direito ilusório".

O artigo pretende responder os trabalhos que contestam frontalmente o direito à explicação e propõe um olhar detido sobre os arts. 13 a 15 do Regulamento, que preveem o direito à "informação significativa sobre a lógica envolvida" em decisões automatizadas, o que, para os autores, sinaliza a existência de um direito à explicação, que deve ser interpretado "funcionalmente e flexivelmente".

Os autores defendem que, ainda que a interpretação do que seria "informação significativa" dependa de uma construção ao longo do tempo, algumas observações já podem ser feitas em relação à sua aplicabilidade. A primeira delas é que este termo se relaciona aos titulares de dados, isto é, que a informação deve ser

[266] SELBST, A. D.; POWLES, J. Meaningful Information and the Right to Explanation. *International Data Privacy Law*, v. 7, n. 4, p. 233-242, 2017. Disponível em: https://papers.ssrn.com/abstract=3039125. Acesso em: 27 maio 2020.

significativa/relevante *para o titular,* alguém que se supõe não ser um *expert* em tecnologia e computação. Em segundo lugar, os autores sustentam que esta análise deve ser funcional, isto é, deve observar o que a explicação em questão agrega à capacidade do titular de se manifestar contra uma decisão, o que seria, para eles, uma abordagem mais relevante do que a que confere um valor intrínseco a esse direito, de difícil mensuração.

Em terceiro lugar, o artigo pontua que deve haver um patamar mínimo para esta funcionalidade ("informação significativa") — a informação prestada deve ser, no mínimo, suficiente para facilitar o exercício, pelo titular, de direitos garantidos pela GDPR e pelo conjunto do Direito Internacional dos Direitos Humanos. Por fim, sustentam que deve haver uma certa flexibilidade quanto ao conteúdo a ser exigido da explicação.

Isso, porque, segundo os autores, a rigidez nas exigências quanto ao que *deveria* corresponder ao direito à explicação é o que teria levado outros autores a defenderem sua inexistência. Um fator que corroboraria o argumento por este modelo flexível de interpretação, focado no aspecto da funcionalidade, é que as diferentes traduções do Regulamento na Europa usam termos correlatos, mas não idênticos, para se referir a "informação significativa".

Feita esta introdução e a defesa da sua tese central, o *paper* passa para uma crítica mais detida da contribuição de Wachter *et al.* ao debate. A primeira crítica dos autores ao argumento sobre a inexistência de um direito à explicação substantivo é que ele parte do pressuposto de que a explicação sobre "decisões específicas" teria uma carga maior do que a explicação sobre "funcionalidade sistêmica". O artigo defende que, pela natureza dos sistemas de *machine learning* ser essencialmente determinística, explicações completas sobre a sua funcionalidade sistêmica deveriam permitir, como consequência, também explicações sobre decisões específicas, de forma que tal distinção, que está no cerne do argumento de Wachter *et al.*, perderia seu sentido.

Uma vez que se aceite o significado funcional de "informação significativa", os autores defendem que é impossível visualizar uma informação significativa, para um titular de dados, sobre a funcionalidade de um sistema de decisões automatizadas, que não seja capaz de oferecer uma explicação sobre decisões específicas. A única possibilidade em que essa separação se sustentaria é se fosse considerada uma abordagem intrínseca de "informação significativa", segundo a qual é suficiente a

informação se ela “dá a sensação ao titular de que ele não está totalmente alijado do controle”. Dessa forma, explicações simples e genéricas sobre a funcionalidade de um sistema, sem entrar nas complexidades que permitem o entendimento de uma decisão específica, poderiam ser consideradas suficientes. Mas, Selbst e Powles defendem que essa lógica contraria o espírito da GDPR de conferir direitos mais robustos aos titulares de dados pessoais.

Quanto à distinção de *timing* estabelecida por Wachter *et al.*, os autores também são categóricos em afastá-la, mais uma vez pela própria natureza dos sistemas de *machine learning.* Além disso, ressaltam que, enquanto o argumento contrário baseia-se em uma análise literal e rígida do texto do Regulamento, ele parece ignorar que os arts. 13 a 15 não fazem nenhuma referência explícita a *timing* de decisões, sendo a única referência a esta distinção proveniente do Considerando 71, que, por sua vez, está relacionado ao art. 22. Considerados estes pontos, os autores afirmam que todos os argumentos de Wachter *et al.* não invalidam a existência de um direito à explicação em si, mas apenas da sua própria versão limitada deste direito.

Selbst e Powles defendem, então, que a posição contrária à sua é falha por se prender a um *framework* analítico excessivamente estático e negligenciar o significado mais profundo das previsões que *de fato* foram incorporadas ao texto final do Regulamento, em especial o significado de “informação significativa”. Também não se debruçam sobre outros dispositivos do Regulamento que reforçariam a existência de um direito à explicação, como o art. 22(3).

Quanto aos argumentos de Wachter *et al.* em relação às limitações apresentadas pelos interesses dos controladores e segredos de negócio/propriedade intelectual, o artigo desconsidera os precedentes citados para sustentar este argumento por serem interpretações da Diretiva, não mais vigente, produzidas no âmbito dos Estados-Membros. Além da mudança na lógica de estabelecimento de precedentes para um modelo mais centralizado com o advento da GDPR[267], Selbst e

[267] Para compreender o que os autores querem dizer quando se referem a uma mudança na lógica de estabelecimento de precedentes de um modelo descentralizado para um modelo mais centralizado com o advento da GDPR, faz-se necessária uma breve nota acerca das especificidades do arranjo jurídico-institucional europeu. De acordo com o art. 288 do Tratado de Funcionamento da União Europeia (TFUE), dispositivo que define os vários tipos de atos jurídicos que a UE pode adotar, as diretivas caracterizam-se como atos jurídicos que vinculam o Estado-Membro destinatário quanto ao resultado a ser alcançado, deixando, no entanto, às instâncias nacionais a competência quanto à forma e aos meios a serem utilizados. Sendo assim, cabe a cada Estado-Membro transpor para o ordenamento

Powles esclarecem que o próprio texto do Regulamento desloca o equilíbrio em favor da explicação. Além disso, também criticam que, ao mesmo tempo em que os autores rejeitam o Considerando 71 para fins de definição de um direito à explicação, recorrem aos Considerandos 47 e 63 para argumentar no sentido dos interesses do controlador e do mercado.

A partir desta cisão fundamental na academia em torno do direito à explicação, outras contribuições foram desenvolvidas de forma a transbordar o debate sobre a existência ou não do direito e cobrir outros aspectos correlatos, alguns que já vinham sendo desenvolvidos no debate mais amplo sobre transparência e *accountability.*

Podemos concluir que a GDPR apresenta uma série de princípios e direitos que implicam, para alguns casos, em um direito à explicação. Esse direito encontra algumas limitações e depende do desenvolvimento de interpretações, principalmente nos órgãos judiciais superiores da União Europeia. Conforme abordaremos nas seções a seguir, a análise de documentos de órgãos regulatórios, decisões de autoridades de proteção de dados ou decisões judiciais no âmbito dos Estados-Membros ajudam a corroborar a tese da existência desse direito e os seus limites.

3.1.2.2 Autoridades de proteção de dados

A análise de alguns precedentes nos tribunais europeus, conforme abordaremos no tópico seguinte, demonstra como a GDPR estabelece um regime de *accountability* e governança algorítmica. Podemos notar, contudo, como a legislação ainda carece de definições precisas. Veremos a seguir como a União Europeia aponta para o desenvolvimento desses temas do ponto de vista regulatório. As autoridades de proteção de dados, contudo, já tem desenvolvido orientações nos marcos da regulação atual que permitem compreender o direito à explicação, orientadas pelo regulamento e pelas recomendações do WP29, agora European Data Protection Board ("EDPB"), sobre decisões automáticas.

Em relatório de 2017, o *Information Commisioner´s Office* (ICO), autoridade de

---

interno as disposições da diretiva na forma de atos legislativos nacionais. Os regulamentos, por sua vez, são atos jurídicos de caráter geral, sendo obrigatórios em todos os seus elementos e diretamente aplicáveis a todos os Estados-Membros. É nesse sentido que se afirma que houve um movimento de centralização com a mudança de regime da Diretiva 95/46/CE para o Regulamento (UE) 679/2016. Cf. UNIÃO EUROPEIA. Tratado sobre o Funcionamento da União Europeia. 6 jul. 2016. Disponível em: https://eur-lex.europa.eu/resource.html?uri=cellar:9e8d52e1-2c70-11e6-b497 01aa75ed71a1.0019.01/DOC_3&format=PDF. Acesso em: 25 nov. 2020.

proteção de dados do Reino Unido, reconhece o direito à explicação para decisões automatizadas ("*right to explanation of a decision based on automated processing*")[268]. Em relatório, destaca que embora argumentem que tal direito só existe nas situações em que apenas processamento automático foi utilizado, essa obrigação deve ocorrer em muitos casos, como nas análises de crédito, no recrutamento e nos seguros, a título de exemplo. Além disso, o documento declara que a *accountability* se tornará um requisito obrigatório sob o regime da GDPR e que, independentemente da organização ou da opacidade inerente às técnicas de inteligência artificial utilizadas, não compreensíveis para seres humanos, seus controladores deverão ter cautela e tomar medidas para garantir informações válidas aos titulares, bem como garantir a justiça, precisão e correção dessas decisões.

Em documento de 2020, a autoridade, em parceria com o The Alan Turing Institute, o instituto nacional de dados e inteligência artificial do Reino Unido, desenvolveu um guia para o desenvolvimento de explicação para inteligência artificial, seja das decisões específicas, ou, em casos em que não seja possível, sobre o funcionamento, justiça e confiabilidade do sistema[269]. Em outro guia publicado dois meses depois, a autoridade traz novas recomendações para o uso de algoritmos e de como as organizações podem explicar seu processamento em um *Data Protection Impact Assessment* (DPIA)[270].

Embora os guias sejam entendidos como recomendações para o desenvolvimento de algoritmos e de serviços, tais instrumentos referem-se aos direitos e as obrigações previstos no regulamento e podem constituir elementos importantes em interpretações judiciais sobre o direito à explicação inserido no regime de *accountability* sistêmica previsto na legislação. O próprio caráter principiológico da GDPR representa uma abertura para interpretações mais amplas sobre os direitos, conforme novas tecnologias e aplicações surgirem. Autores importantes têm reconhecido as limitações do princípio da transparência para garantir todos os direitos

---

[268] INFORMATION COMMISSIONER'S OFFICE. *Big data, artificial intelligence, machine learning and data protection*. 2017. Disponível em: https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf. Acesso em: 28 jul. 2020.

[269] INFORMATION COMMISSIONER'S OFFICE; THE ALAN TURING INSTITUTE. ICO and Turing consultation on Explaining AI decisions guidance. *ICO*, 24 jan. 2020. Disponível em: https://ico.org.uk/about-the-ico/ico-and-stakeholder-consultations/ico-and-the-turing-consultation-on-explaining-ai-decisions-guidance/. Acesso em: 16 abr. 2021.

[270] INFORMATION COMMISSIONER'S OFFICE. *Guidance on the AI auditing framework*. [*S. l.*]: ICO, 2020. Disponível em: https://ico.org.uk/media/about-the-ico/consultations/2617219/guidance-on-the-ai-auditing-framework-draft-for-consultation.pdf. Acesso em: 1 ago. 2020.

dos titulares.

A GDPR é um marco inserido numa estratégia da União Europeia para as transformações digitais. Os princípios e direitos previstos no regulamento visam garantir um ambiente confiável e robusto para o desenvolvimento econômico associado à economia digital e principalmente à inteligência artificial. A Comissão Europeia, órgão responsável por aplicar as decisões do Parlamento Europeu, bem como propor legislações e medidas de defesa dos interesses da União Europeia, tem desenvolvido uma série de documentos que devem orientar os desenvolvimentos regulatórios no bloco e nos países membros. Em 2018, a Comissão Europeia, no documento "Coordinated plan on Artificial Intelligence", compara o surgimento da inteligência artificial ao desenvolvimento da eletricidade na capacidade de transformação da sociedade[271].

Os desafios associados a essa transformação dizem respeito ao medo dos trabalhadores de perderem seu emprego, aos riscos aos consumidores em relação à responsabilidade por erros causados por falhas nos algoritmos ou mesmo às futuras máquinas operadas por inteligência artificial. As pequenas empresas temem não conseguirem aplicar a inteligência artificial nos seus negócios de forma adequada e as *startups* temem não conseguirem competir com as concorrentes dos EUA e da China. Nesse plano coordenado estão presentes alguns objetivos principais: a criação de um *pool* de dados englobando todo o bloco, o incentivo à criação de talentos e conhecimento na área bem como o fortalecimento da confiança na inteligência artificial.

A implementação dessa estratégia depende, segundo a comissão, do desenvolvimento de conhecimento avançado em técnicas de inteligência artificial, uma quantidade massiva de dados e de um corpo regulatório robusto e eficiente que permita a criação de um ambiente de excelência e confiança[272]. Esse ambiente inclui a garantia dos valores da UE e da proteção de dados pessoais como a transparência, a capacidade de os cidadãos interagirem com os sistemas e entenderem seu funcionamento.

---

[271] UNIÃO EUROPEIA. *COM (2018) 785 — Coordinated Plan for Artificial Intelligence*, 2018. Disponível em: https://ec.europa.eu/knowledge4policy/publication/coordinated-plan-artificial-intelligence-com2018-795-final_en#:~:text=7%20December%202018-,Coordinated%20Plan%20on%20Artificial%20Intelligence%20. Acesso em 13 jul. 2020.

[272] UNIÃO EUROPEIA. *COM (2019) 168 — Building Trust in Human-Centric Artificial Intelligence*. 2019. Disponível em: https://ec.europa.eu/transparency/regdoc/rep/1/2019/EN/COM-2019-168-F1-EN-MAIN-PART-1.PDF. Acesso em: 13 jul. 2020.

Esses documentos têm afirmado que, para garantia dos valores da União Europeia e dos direitos dos cidadãos, as novas tecnologias devem investir, desde o *design*, em mecanismos que garantam a governança, a segurança e *accountability*. Para garantir os direitos, entre as metas, a Comissão recomenda o desenvolvimento da explicabilidade dos sistemas[273].

Em 2019, um relatório do HLEG (*High Level Expert Group in Artificial Intelligence*) encomendado pela Comissão criou um guia ético para o desenvolvimento de inteligência artificial[274]. O objetivo do relatório é apresentar um guia de ética no desenvolvimento e uso de inteligência artificial como forma de garantir o desenvolvimento de uma inteligência artificial confiável (*trustworthy artificial intelligence*), correspondendo à meta europeia de estabelecer um ambiente propício para o desenvolvimento econômico e tecnológico de forma a respeitar os direitos fundamentais. O guia foca o aspecto ético e técnico para atingir a inteligência artificial confiável.

De acordo com o guia, a inteligência artificial confiável deve possuir três elementos, que deverão estar presentes em todo o ciclo de vida do sistema: a) legalidade (*lawfulness*); b) ética e; c) robustez sob o ponto de vista técnico e social. O guia oferece elementos para lidar com os dois últimos e argumenta que muitas vezes as recomendações se confundem com os deveres legais, visto que se baseiam nos direitos fundamentais.

Os direitos fundamentais que servem de parâmetro no guia são a dignidade da pessoa humana, a liberdade individual, a democracia, o devido processo legal, a igualdade, a não discriminação e os direitos dos cidadãos. A maioria dos direitos fundamentais estão positivados na legislação. O guia, contudo, delimita alguns princípios necessários à garantia desses direitos no contexto da inteligência artificial. Isso permite pensar em termos do que se deve fazer do ponto de vista ético e técnico e não apenas no que é permitido fazer em termos legais.

Os princípios éticos devem guiar o desenvolvimento e aplicação de inteligência artificial em todas as fases. Através deles é possível encontrar soluções no

---

[273] UNIÃO EUROPEIA. *COM (2018) 237 — Artificial Intelligence for Europe*. 2018. Disponível em: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A237%3AFIN. Acesso em: 13 jul. 2020.

[274] Idem. *Ethics Guidelines for Trustworthy Artificial Intelligence*. European Commission, 2019. Disponível em: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai. Acesso em: 13 jul. 2018.

desenvolvimento e fazer com que sejam éticas além de meramente adequadas a legislações. Em casos de conflito de princípios, pode-se sopesar e encontrar as melhores soluções, embora em algumas condições, alguns direitos nunca possam ser relativizados. Esses princípios são o respeito à autonomia humana, a prevenção de danos, a justiça (*fairness*) e a explicabilidade.

Diante desses princípios, o guia então recomenda alguns requisitos para a implementação da inteligência artificial. Os desenvolvedores devem aplicar esses requisitos desde o desenho, os aplicadores devem buscar conhecer se os sistemas que atendem a esses requisitos e os usuários devem ser informados dos requisitos e ter mecanismos para verificar seu cumprimento. São eles: a) agência e supervisão humana, que inclui mecanismos de governança como HITL (*human-in-the-loop*), HOTL (*human-on-the-loop*) ou HIC (*Human in comand*); b) robustez técnica e segurança necessária para atingir o princípio da prevenção de danos que envolve resiliência contra ataques, planos de contingência, precisão nas decisões tomadas, garantida por meio de processos adequados para garantir a correção das decisões erradas, confiabilidade e reprodutibilidade; c) privacidade e governança de dados relacionado ao princípio da prevenção de dados, o que envolve o desenvolvimento de sistemas respeitando a privacidade e a proteção de dados; d) transparência relacionada ao princípio da explicabilidade deve estar presente nos dados, no sistema e no modelo de negócios; e) diversidade e não discriminação; f) bem estar social e ambiental e g) *accountability*.

Após um período de discussões e apontamentos, em fevereiro de 2020 a Comissão lançou outro documento importante, o "*White Paper: On Artificial Intelligence — A European approach to excellence and trust"*[275] . O documento faz parte de uma série de avaliações sobre a implementação da GDPR e teve o objetivo de apresentar diretrizes para o desenvolvimento de um ambiente de excelência, confiança e respeito aos direitos fundamentais na Europa. As recomendações têm o objetivo de evitar a fragmentação regulatória e avançar na consolidação de um mercado único de dados. As discussões regulatórias visam garantir que a UE seja uma referência no desenvolvimento de tecnologias na nova onda de dados que se inicia.

[275] UNIÃO EUROPEIA. *COM (2020) 65 — White Paper on Artificial Intelligence: a European approach to excellence and trust*. Disponível em: https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en. Acesso em: 01 ago. 2020.

Além disso, em relação ao tema desse trabalho, podemos notar como a explicabilidade ocupa um lugar central no desenvolvimento do ambiente de confiança para a inteligência artificial. O documento resgata os apontamentos do relatório do HLEG e argumenta que dos pontos apontados pelo relatório, os que mais demandam atenção dizem respeito à transparência, à rastreabilidade e à supervisão humana.

Podemos notar como a explicabilidade constitui um elemento central no regime do regulamento e de como podemos notar a existência de um direito subjetivo à explicação, ao menos em determinados casos previstos na lei. Em outros casos, além disso, nos quais a explicação se constitua como necessária para o exercício de direitos fundamentais no contexto da proteção de dados, esse direito também pode ser invocado a partir da interpretação sistemática e teleológica do Regulamento e da legislação da UE.

3.1.2.3 As interpretações dos tribunais

O entendimento jurisprudencial sobre privacidade e proteção de dados vem se desenvolvendo, na Europa, desde pelo menos a década de 1980, acompanhando a evolução da legislação. Os principais marcos legais da Europa, que promoveram para outros países um modelo de proteção de dados, foram a Convenção 108 do Conselho da Europa e a Diretiva 95/46/CE, de 1995[276], até a aprovação do regulamento em vigor.

O texto da Diretiva já apresentava um modelo de regulação mais forte e mais amplo, e as opiniões do WP29 apresentavam o quadro que culminaria no Regulamento. Apesar da evidente continuidade nos conceitos e na intencionalidade desses documentos, há uma forte mudança, principalmente em relação à força do regulamento, ao vincular diretamente a conduta de todos os agentes da União Europeia.

O Regulamento deixa estabelecida a obrigação de notificação de decisões automatizadas, inclusive *profiling*, por parte dos controladores de dados pessoais e o direito à informação para os titulares. Ele ainda cria procedimentos e direitos relacionados ao uso de algoritmos. Kaminski argumenta que o Regulamento, interpretado a partir das Considerandas, em relação ao § 3º do art. 22, cria um regime

[276] GREENLEAF, G. The Influence of European Data Privacy Standards Outside Europe: Implications for Globalisation of Convention 108. *International Data Privacy Law*, [*S. l.*], v. 2, 2011.

de *accountability* algorítmica[277]. No entanto, sobre a aplicação desse regime, o texto deixa bastante espaço para interpretação.

As interpretações possíveis relacionam-se até mesmo a outros dispositivos do Regulamento, como o conceito de dado pessoal. Wachter e Mittelstadt discutem as interpretações das cortes europeias sobre possibilidade de inferências serem encaradas como dados pessoais[278]. As decisões anteriores ao Regulamento, segundo os autores, nem sempre seguem a concepção mais expansionista de dados pessoais que se pode deduzir do texto deste diploma.

Em uma das aplicações mais utilizadas de algoritmos e inteligência artificial, o *profiling*, para fins de marketing, produz apenas inferências probabilísticas que não necessariamente são vinculadas a dados fornecidos pelo titular. Nesse sentido, a possibilidade de retificação das informações pode não ser suficiente para garantir a autodeterminação dos usuários. Wachter e Mittelstadt argumentam então que é necessário garantir o direito a inferências razoáveis como forma de evitar essa lacuna.

Outra barreira importante que ainda deixa um largo campo para as interpretações diz respeito à propriedade intelectual e o segredo de negócio em relação à transparência como um todo. Mesmo as informações que devem ser fornecidas pelos titulares sobre a existência de decisões automatizadas podem encontrar um limite no direito à propriedade intelectual.

Outros pontos, que ainda serão melhor definidos pela jurisprudência e pela atuação das autoridades de proteção de dados, dizem respeito aos "efeitos significativos" presentes no art. 22 como condição para não se submeter o titular a decisões automatizadas. Outra questão, ainda sobre o mesmo artigo, apontada por Wachter *et al.*[279], diz respeito à previsão de o direito não se submeter a decisões baseadas apenas em processamento automatizado. Interpretada de maneira literal, a menor intervenção humana pode limitar a aplicabilidade do requerimento de informações significativas ou mesmo o direito a não se submeter a tais decisões.

---

[277] KAMINSKI, M. E. The Right to Explanation, Explained. *Berkeley Technology Law Journal*, v. 34, n. 189, 2019. Disponível em: https://papers.ssrn.com/abstract=3196985. Acesso em: 27 maio. 2020.

[278] WACHTER, S.; MITTELSTADT, B. A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI. *LawArXiv*, 12 out. 2018. Disponível em: https://doi.org/10.31228/osf.io/mu2kf. Acesso em: 27 maio 2020.

[279] WACHTER, S.; MITTELSTADT, B.; FLORIDI, L. *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation*. Rochester, NY: Social Science Research Network, 2016. Disponível em: https://papers.ssrn.com/abstract=2903469. Acesso em: 27 maio. 2020.

Selbst e Powles[280] argumentam, contudo, que a interpretação das autoridades de proteção de dados deverá ser no sentido de entender que envolvimentos triviais de seres humanos não serão suficientes para descaracterizar determinado processamento como sendo automatizado.

Algumas dessas questões já são enfrentadas pelos tribunais nacionais e pelas autoridades reguladoras. Na Holanda, a justiça proibiu o uso do Siri, sistema desenvolvido por órgãos do governo para detecção de fraude na seguridade social. O sistema classificava, com base em uma série de informações de bancos de dados públicos, o risco de fraude para cada cidadão atendido por programas governamentais em determinada vizinhança[281]. O sistema foi muito criticado por entidades da sociedade civil que questionaram sua utilização e a possibilidade de que fosse utilizado para aumentar a vigilância da população vulnerável. Além disso, argumentou que essa classificação era feita de forma automatizada e sem notificação aos titulares dos dados. A sentença da corte de Haia se apresenta como um caso interessante para análise das questões de *accountability* algorítmica [282].

O governo argumentava, em sua defesa, que os dados eram utilizados por meios legítimos e para fins legítimos. Além disso, que a classificação em si não acarretava nenhum efeito legal, que servia apenas como indicativo para futuras investigações de funcionários do governo e que, portanto, não estaria sujeita à definição de decisão automatizada do art. 22. Argumentou, ainda, que a revelação de informações relacionadas ao funcionamento do sistema possibilitaria que fraudadores tomassem medidas para fugir dessa avaliação de risco.

A decisão considerou que a mera classificação de risco causa impactos nos cidadãos classificados e que o governo não ofereceu salvaguardas suficientes para garantir que o seu uso não poderia ter um resultado discriminatório sobre determinadas regiões ou grupos de pessoas.

A corte, contudo, não enfrentou a questão de que a mera análise de risco de fraude, em si, se enquadraria na definição ou nas exceções previstas no art. 22 do

---

[280] SELBST, A. D.; POWLES, J. Meaningful Information and the Right to Explanation. *International Data Privacy Law*, v. 7, n. 4, p. 233-242, 2017. Disponível em: https://papers.ssrn.com/abstract=3039125. Acesso em: 27 maio 2020.

[281] BRAUN, I. *High-Risk Citizens*. 2018. Disponível em: https://algorithmwatch.org/en/story/high-risk-citizens/. Acesso em: 28 jul. 2020.

[282] RB. DEN HAAG. *ECLI:NL:RBDHA:2020:1878, Rechtbank Den Haag, C-09-550982-HA ZA 18-388*. 2020. Disponível em: https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBDHA:2020:1878. Acesso em: 28 jul. 2020.

Regulamento, mas entendeu que o *software* violava a privacidade dos cidadãos afetados e determinou que os órgãos do governo interrompessem a sua utilização. O principal argumento do tribunal para a proibição do algoritmo foi a falta de adequação do tratamento de dados e a desproporcionalidade das informações utilizadas.

Um ponto de destaque sobre a decisão diz respeito ao juízo sobre elementos relacionados a *accountability* dos algoritmos. A corte considerou a falta de transparência sobre as informações utilizadas, sobre a lógica do sistema, sobre a impossibilidade de se avaliar a integridade das decisões e a confiabilidade das previsões do algoritmo. Além disso, que os cidadãos, cujos dados foram processados pelo sistema, não sabiam da existência do processamento e tampouco tinham condições de questionar seu resultado.

Um outro caso, julgado na mesma corte holandesa, enfrenta questões parecidas sobre *accountability* algorítmica, no entanto, em uma aplicação que foi considerada adequada[283]. O Ministério da Justiça e Segurança da Holanda, no contexto da avaliação do potencial de risco para autorização do uso de armas de caça, desenvolveu, em parceria com pesquisadores da área de psiquiatria e psicologia vinculados a universidades e institutos de pesquisa, um questionário eletrônico que avalia fatores de risco para porte de armas.

O *e-Screener* é composto por 99 questões que avaliam 10 fatores de risco relacionados a características psicológicas dos requerentes à autorização do porte de armas. Ele foi desenvolvido com base em uma amostra de pessoas de alto risco e de pessoas de baixo risco com porte de armas. O questionário foi então avaliado e calibrado por outro instituto, que produziu um relatório sobre o sistema com algumas recomendações de melhoria.

Em 2019 o sistema foi posto em prática e diversos novos requerentes, bem como antigos detentores da licença, submeteram-se aos testes. Os testes negativos não significam uma rejeição automática da licença, no entanto, dependem de justificativa do requerente para que tenha a licença concedida. O resultado do sistema, bem como outros elementos, é analisado pelos delegados de polícia locais, que podem fornecer ou não a autorização para uso de armas de caça.

---

[283] RB. DEN HAAG. *ECLI:NL:RBDHA:2020:1013*, *Rechtbank Den Haag, C-09-585239-KG ZA 19-1221*. 2020. Disponível em: https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBDHA:2020:1013. Acesso em: 28 jul. 2020.

Uma associação de caçadores questionou o uso do *e-Screener*, argumentando, entre outras coisas, que a confiabilidade do sistema não foi demonstrada, que ele viola o direito ao contraditório, que não garante condições para pessoas com deficiência e que viola a GPDR. Argumenta, além disso, que o resultado é utilizado como análise final e não como mero auxílio na decisão. A associação solicitou ainda a divulgação de documentos relacionados ao desenvolvimento e implementação do sistema, incluindo o Ministério da Justiça e Segurança holandês, as entidades desenvolvedoras e o instituto que produziu o relatório de avaliação do *software*. No pedido da associação de caçadores, todos os documentos relacionados a alguns temas deveriam ser revelados: os fundamentos científicos e teóricos, os dados sobre a confiabilidade psicométrica e sobre a determinação da pontuação de corte, investigações sobre a justiça do programa para determinados grupos sociais, como jovens ou idosos, e as perguntas utilizadas no teste.

A associação argumentou que não haveria comprovação científica de que determinadas características psicológicas levem ao maior número de incidentes com armas de fogo e que os casos holandeses são tão poucos que não é possível fazer tais inferências. O juiz de primeira instância entendeu que os *experts* consultados realizaram a avaliação de riscos baseando-se em evidências científicas de que determinadas características psicológicas implicam em maior risco.

Outro argumento apresentado contra o uso do sistema dizia respeito à calibração e validação. O juiz entendeu, contudo, que o relatório estava completo e apontava os riscos, falhas e limitações do instrumento, de forma que não se poderia objetar sua implementação.

A demandante argumentou, ainda, que devido ao fato do funcionamento do sistema e dos elementos determinantes das conclusões permanecerem secretos, o uso do *e-Screener* feria a igualdade de condições no processo. Ainda sobre o tema, argumenta que o delegado de polícia, por também não saber sobre o funcionamento do sistema, não seria capaz de avaliar de maneira adequada o seu resultado. A corte entendeu, em acordo com a primeira instância, que durante a implementação do sistema houve um treinamento adequado aos encarregados para sua utilização e que a divulgação dos elementos e da pontuação permitiria que se desenvolvessem métodos para burlar os seus resultados.

A corte entendeu que o governo atendia aos requisitos dos arts. 14 e 15 do regulamento sobre os direitos relacionados à informação aos titulares dos dados e

manteve a decisão de primeira instância sobre a legalidade da utilização do *e-Screener*. É importante notar que a decisão avaliadora dos algoritmos efetuou uma abordagem sistêmica, levando em consideração os cuidados no desenvolvimento da aplicação bem como os meios de questionamento em um devido processo legal.

Vale destacar, ainda, que o direito ao devido processo não implicou na divulgação completa dos códigos de programação, tampouco sobre os critérios para produção dos valores de referência. No entanto, a corte entendeu que o governo forneceu elementos suficientes para demonstrar o cuidado no desenvolvimento do sistema, a correta avaliação de riscos, o treinamento para a correta aplicação do sistema, bem como meios para questionamento dos seus resultados.

Embora as duas aplicações se relacionem a contextos diversos e sejam tipos de algoritmos tecnicamente diferentes, é possível tirar algumas conclusões na comparação entre as duas. A possibilidade de os titulares entenderem a lógica do algoritmo, seu contexto e ao menos as categorias de informações utilizadas para determinadas conclusões foram elementos importantes para a consideração de adequação no seu uso. Além disso, demonstram como o uso de algoritmos inseridos em um devido processo legal tornam sua aplicação menos questionável.

Convém notar, além disso, em consonância com algumas organizações em defesa da proteção de dados[284], como o art. 22 do Regulamento é de difícil aplicação e demanda maior desenvolvimento hermenêutico e dogmático[285], seja por meio de orientações ou por meio de desenvolvimento legislativo. Alguns conceitos, como

---

[284] PRIVACY INTERNATIONAL. Data Is Power: Profiling and Automated Decision-Making in GDPR. [s.d.]. Disponível em: http://privacyinternational.org/report/1718/data-power-profiling-and-automated-decision-making-gdpr. Acesso em: 31 jul. 2020.

[285] “Art 22. Automated individual decision-making, including profiling […] 1. The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her. 2. Paragraph 1 shall not apply if the decision: (a) is necessary for entering into, or performance of, a contract between the data subject and a data controller; (b) is authorised by Union or Member State law to which the controller is subject and which also lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests; or (c) is based on the data subject's explicit consent. 3. In the cases referred to in points (a) and (c) of paragraph 2, the data controller shall implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision. 4. Decisions referred to in paragraph 2 shall not be based on special categories of personal data referred to in Article 9(1), unless point (a) or (g) of Article 9(2) applies and suitable measures to safeguard the data subject's rights and freedoms and legitimate interests are in place.” (UNIÃO EUROPEIA. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance)*. Official Journal of the European Union, 4 maio 2016. Disponível em; https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng. Acesso em: 27 ago. 2021).

“decisões baseadas apenas em processamento automático”, de “efeitos significativos”, ou a intervenção humana são por demais amplos. No caso do Siri, por exemplo, apesar dos efeitos potencialmente nocivos e do *profiling* representar, em si, um possível dano aos titulares, não foi possível à corte definir se tal uso do algoritmo consistia no conceito previsto no artigo.

Os tribunais e as cortes europeias terão o desafio de enfrentar a questão sobre o direito a informações significativas em outros casos, como o trazido à tona por motoristas-parceiros em plataformas digitais de intermediação entre passageiros e motoristas. Essas plataformas podem realizar *profiling* e operacionalizar milhões de decisões automatizadas na distribuição de corridas por meio de seus algoritmos. Protegidos por direitos de propriedade intelectual, essas aplicações estão sofrendo questionamentos judiciais sobre a lógica envolvida, os dados coletados e informações produzidas sobre os titulares.

O sindicato dos trabalhadores em aplicativos, o *App Drivers & Couriers Union* (ADCU), entrou com uma representação, na Holanda, para que a empresa Uber forneça acesso aos dados pessoais e informações sobre a lógica por trás de classificações baseadas no comportamento dos motoristas[286]. O sindicato argumenta que a empresa não cumpre as obrigações de fornecer os dados pessoais sob seu controle além de descumprir as obrigações referentes à tomada de decisões automatizadas.

Não restam dúvidas de que veremos nos próximos anos, em diversas aplicações, o questionamento do uso de algoritmos e de decisões automatizadas, demandando o acesso aos dados pessoais ou informações significativas sobre seu funcionamento. É prudente imaginar como testemunharemos uma importante discussão sobre o direito à explicação, à propriedade intelectual, limitações técnicas ou legais. O entendimento dos tribunais superiores poderá lançar mais clareza sobre essas questões.

No entanto, embora as questões jurisdicionais sejam de suma importância, as opiniões, regulamentos e recomendações das autoridades cumprirão um papel essencial nesse desenvolvimento, visto que também orientarão as decisões judiciais. Nada impede, além disso, que alterações na legislação vigente ocorram para dar

[286] ACDU. Uber Drivers Take Unprecedented International Legal Action To Demand Their Data. *ACDU*, 20 jul. 2020. Disponível em: https://www.adcu.org.uk/news-posts/uber-drivers-take-unprecedented-international-legal-action-to-demand-their-data. Acesso em: 31 jul. 2020.

maior clareza e aplicabilidade nestas questões.

3.2 CENÁRIO NACIONAL

Uma vez apresentado o panorama do debate internacional, passaremos a discorrer mais especificamente sobre os instrumentos legais aptos a fundamentar a existência de um direito à explicação no ordenamento jurídico brasileiro. Como é sabido, a Lei nº 13.709/2018 não surgiu no contexto de um vazio normativo. Com a sua promulgação, a LGPD passou a coexistir com diversas normas setoriais que, direta ou indiretamente, endereçavam e continuam a endereçar a proteção de dados pessoais em setores específicos.

Dentre essas normas, destacamos a legislação versando sobre transparência e direito à informação em matéria consumerista, em especial o Código de Defesa do Consumidor (Lei nº 8.078/1990) e a Lei do Cadastro Positivo (Lei nº 12.414/2011), que fornecem fortes elementos para a fundamentação de um direito à explicação, ainda que restrito às relações de consumo.

Nesse ponto, também buscamos analisar a jurisprudência recente do Superior Tribunal de Justiça em julgamentos envolvendo a legalidade de cadastros negativos e positivos de crédito, bem como a utilização de sistemas de pontuação de crédito. Em conjunto, esses julgados reforçam os direitos de transparência e acesso à informação já estabelecidos nas referidas legislações, e dão ainda um passo além ao reconhecerem a necessidade de um efetivo devido processo informacional no contexto de decisões automatizadas que venham a impactar o livre desenvolvimento da personalidade dos consumidores. Ademais, apesar de versar especificamente sobre direitos de transparência e acesso à informação em relações de consumo, a literatura mais recente em torno desses julgados admite a possibilidade de expansão de seu escopo para situações envolvendo o emprego de sistemas automatizados em outros contextos que não o consumerista.

Em seguida, passamos a analisar de que forma, ou a partir de quais dispositivos, seria possível fundamentar a existência de um direito à explicação com base no texto da LGPD. Nesse sentido, destacamos os objetivos, fundamentos, princípios e direitos que, uma vez combinados e interpretados em seu conjunto, permitem-nos fundamentar a existência de um direito à explicação. Em seguida, numa tentativa de reconstrução do debate brasileiro em torno das decisões automatizadas

e do direito à explicação na LGPD, buscamos destacar as especificidades e insuficiências dessas disposições, que, como veremos, ora permitem reconhecer um direito à explicação de escopo significativamente mais amplo, ora um direito significativamente mais restrito do que aquele derivado do regulamento europeu.

3.2.1 As regulações setoriais nacionais

Um dos setores da economia e do mercado que mais se vale do uso e tratamento de dados pessoais, principalmente para viabilizar decisões automatizadas para ofertar seus serviços, é o de consumo. Esse setor é caracterizado pela necessidade de se entender o consumidor e, inclusive, influenciar seus hábitos. No entanto, neste cenário, o consumidor se encontra em posição vulnerável em sua relação com as empresas[287], e, por isso, deve ser protegido de forma mais robusta. Entre as medidas de proteção, deve-se incluir o fornecimento de informações adequadas para que possa exercer seus direitos e evitar práticas abusivas e discriminatórias.[288] À medida que modelos de negócio se baseiam cada vez mais na coleta e processamento de dados pessoais, o uso intenso desse tipo de informação pode levar a práticas indesejadas, abusivas e prejudiciais (conforme ilustrado nos exemplos no início deste trabalho). Para entender como o mercado de consumo instrumentalizou ferramentas para combater tais práticas, analisamos o Código de Defesa do Consumidor e a Lei do Cadastro Positivo, duas regulações setoriais que formam um robusto ecossistema de proteção de dados.

O Código de Defesa do Consumidor (CDC), Lei nº 8.078/1990, é uma regulação setorial que se aplica às relações de consumo, sejam elas *on-line* ou *off-line*, e

---

[287] "Art. 4º. A Política Nacional das Relações de Consumo tem por objetivo o atendimento das necessidades dos consumidores, o respeito à sua dignidade, saúde e segurança, a proteção de seus interesses econômicos, a melhoria da sua qualidade de vida, bem como a transparência e harmonia das relações de consumo, atendidos os seguintes princípios: I — reconhecimento da vulnerabilidade do consumidor no mercado de consumo." (BRASIL. *Lei nº 8.078, de 11 de setembro de 1990*. Dispõe sobre a proteção do consumidor e dá outras providências. Brasília, DF: Presidência da República, 12 set. 1990. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l8078compilado.htm. Acesso em: 8 abr. 2021).

[288] "Art. 6º. São direitos básicos do consumidor: [...] III — a informação adequada e clara sobre os diferentes produtos e serviços, com especificação correta de quantidade, características, composição, qualidade, tributos incidentes e preço, bem como sobre os riscos que apresentem; [...] IV — a proteção contra a publicidade enganosa e abusiva, métodos comerciais coercitivos ou desleais, bem como contra práticas e cláusulas abusivas ou impostas no fornecimento de produtos e serviços. (BRASIL. *Lei nº 8.078, de 11 de setembro de 1990*. Dispõe sobre a proteção do consumidor e dá outras providências. Brasília, DF: Presidência da República, 12 set. 1990. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l8078compilado.htm. Acesso em: 8 abr. 2021).

estabelece a transparência e a boa-fé como princípios que orientam essas relações.[289] No que diz respeito à boa fé em sua forma objetiva,[290] é entendimento do Superior Tribunal de Justiça que ela "[...] seria como um modelo ideal de conduta, que se exige de todos os integrantes da relação obrigacional (devedor e credor) na busca do correto adimplemento da obrigação, que é a sua finalidade."[291] A interpretação conjunta do CDC e da decisão do STJ aponta para o dever de informar o consumidor de maneira clara e objetiva a respeito da relação contratual, o que inclui o período pré-negocial, e o dever de máxima transparência dos arquivos de consumo. Nesse sentido, o dever de informação se deve às obrigações derivadas da boa-fé objetiva.

Destacam-se dois artigos do CDC que tratam do acesso a informações cadastrais e em bancos de dados. O primeiro deles, o art. 43, ao regular os arquivos de consumo, deixou expresso o direito de acesso do consumidor, nesses cadastros e bancos de dados, a informações a seu respeito e às respectivas fontes. Também determinou o dever de clareza dos arquivos, o direito de retificação de informações incorretas e que o consumidor deve ser notificado[292] sobre a coleta e o uso de seus dados, ainda que o consentimento prévio não seja necessário — com a exceção de casos de compartilhamento com terceiros, conforme o entendimento do Ministério da Justiça.[293] Além disso, estipula um período máximo de armazenamento dos dados do

---

[289] Ver art. 4º do Código de Defesa do Consumidor."

[290] A boa-fé subjetiva se refere ao estado psicológico da pessoa, consistente na justiça, ou, na licitude de seus atos, ou na ignorância de sua antijuridicidade. Já a boa-fé objetiva consiste em um dever ativo de conduta contratual de ambos os contratantes e os obriga a colaborar e cooperar, levando em consideração os interesses um do outro, a fim de alcançar o efeito prático que justifica a existência jurídica do contrato celebrado. Cf. COELHO, F. U. *Curso de Direito Comercial.* Direito de Empresa. São Paulo: Revista dos Tribunais, 2018. 2 v.

[291] BRASIL. Superior Tribunal de Justiça. *Recurso Especial nº 1.200.105 — AM 2010/0111335-0*. Relator: Min. Paulo de Tarso Sanseverino. Data de Julgamento: 21/05/2013. Data de Publicação: 27/05/2013. Disponível em: https://stj.jusbrasil.com.br/jurisprudencia/23342165/embargos-de-declaracao-no-recurso-especial-edcl-no-resp-1200105-am-2010-0111335-0-stj/inteiro-teor-23342166?ref=amp. Acesso em: 17 ago. 2020.

[292] "Art. 43. O consumidor, sem prejuízo do disposto no art. 86, terá acesso às informações existentes em cadastros, fichas, registros e dados pessoais e de consumo arquivados sobre ele, bem como sobre as suas respectivas fontes. § 1° Os cadastros e dados de consumidores devem ser objetivos, claros, verdadeiros e em linguagem de fácil compreensão, não podendo conter informações negativas referentes a período superior a cinco anos. § 2° A abertura de cadastro, ficha, registro e dados pessoais e de consumo deverá ser comunicada por escrito ao consumidor, quando não solicitada por ele. § 3° O consumidor, sempre que encontrar inexatidão nos seus dados e cadastros, poderá exigir sua imediata correção, devendo o arquivista, no prazo de cinco dias úteis, comunicar a alteração aos eventuais destinatários das informações incorretas." (BRASIL. *Lei nº 8.078, de 11 de setembro de 1990*. Dispõe sobre a proteção do consumidor e dá outras providências. Brasília, DF: Presidência da República, 12 set. 1990. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l8078compilado.htm. Acesso em: 8 abr. 2021).

[293] BRASIL. Ministério da Justiça. *Portaria nº 5 de 27 de agosto de 2002*. Dispõe sobre cláusulas abusivas em contratos de vendas de produtos e prestação de serviços. Diário Oficial da República

consumidor de 5 anos. Já o art. 46 determina que

> Art. 46. Os contratos que regulam as relações de consumo não obrigarão os consumidores, se não lhes for dada a oportunidade de tomar conhecimento prévio de seu conteúdo, ou se os respectivos instrumentos forem redigidos de modo a dificultar a compreensão de seu sentido e alcance.[294]

O artigo não só reafirma o direito à informação sobre a relação de consumo, mas também determina que deve ser repassada de forma inteligível, para garantir a sua compreensão.

Dessa forma, quando houver decisão automatizada no contexto de uma relação de consumo, como a concessão ou não de um financiamento de veículo, por exemplo, o consumidor tem o direito de ter acesso aos (seus) dados que basearam a tomada de decisão. Caso seja criada uma obrigação jurídica, é seu direito, também, ter conhecimento do sentido desta, ou seja, as suas finalidades e propósitos, seu alcance e como ela foi formada, incluindo critérios e valoração dos atributos utilizados para tomar a decisão. Em outras palavras, entender como se deu a formação da obrigação jurídica é essencial para a sua aceitação e exercício dos direitos previstos no CDC. E isso incluiria entender como um algoritmo deu origem a tal obrigação.

Esta lógica também foi empregada pela Lei do Cadastro Positivo (Lei nº 12.414/2011, LCP), atualizada pela Lei Complementar nº 166/2019[295], que estabelece normas voltadas à "[...] disciplina e consulta a bancos de dados com informações de adimplemento, de pessoas naturais ou de pessoas jurídicas, para a formação de histórico de crédito." Entre os principais objetivos desta lei estão a redução da assimetria de informações e possibilitar a coleta de dados de adimplência após o consentimento prévio do consumidor. Afirma-se que isso possibilitaria a redução de taxas de juros e uma consequente ampliação das relações comerciais, o que favoreceria e protegeria todo o ecossistema consumerista. A norma visa, também, a

---

Federativa do Brasil, Brasília-DF, 28 ago. 2002. Disponível em: https://www.procon.go.gov.br/legislacao/portarias/portaria-n%C2%BA-5-27-08-2002-mj-sde-clausulas-abusivas-nome-de-consumidor-a-banco-dedados.html. Acesso em: 20 out. 2018

[294] BRASIL. *Lei nº 8.078, de 11 de setembro de 1990*. Dispõe sobre a proteção do consumidor e dá outras providências. Brasília, DF: Presidência da República, 12 set. 1990. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l8078compilado.htm. Acesso em: 8 abr. 2021.

[295] BRASIL. *Lei Complementar nº 166, de 8 de abril de 2019*. Altera a Lei Complementar nº 105, de 10 de janeiro de 2001, e a Lei nº 12.414, de 9 de junho de 2011, para dispor sobre os cadastros positivos de crédito e regular a responsabilidade civil dos operadores. Brasília, DF: Presidência da República, 08 abr. 2019. Disponível em: http://www.planalto.gov.br/ccivil_03/LEIS/LCP/Lcp166.htm#art2. Acesso em: 8 abr. 2021.

adequada proteção de dados pessoais de consumo, ao prever uma série de novos direitos, entre eles alguns que favorecem o argumento da existência de um direito à explicação. Nesse contexto, destacam-se os direitos previstos no art. 5º, tais como os de: "IV — conhecer os principais elementos e critérios considerados para a análise de risco, resguardado o segredo empresarial"; "V — ser informado previamente sobre o armazenamento, a identidade do gestor do banco de dados, o objetivo do tratamento dos dados pessoais e os destinatários dos dados em caso de compartilhamento"; "VI — solicitar ao consulente a revisão de decisão realizada exclusivamente por meios automatizados"; e "VII — ter os seus dados pessoais utilizados somente de acordo com a finalidade para a qual eles foram coletados."[296]

Estes quatro direitos se originam a partir do direito à transparência e não discriminação e formam a espinha dorsal do direito à explicação de decisões automatizadas em relações de consumo. Eles exigem que o consumidor seja esclarecido sobre as fontes de dados utilizadas e as informações pessoais consideradas para o cálculo do risco de inadimplência na concessão ou não de crédito. A lei também busca limitar os tipos de dados que podem ser utilizados para cálculo do risco de crédito, vedando o uso de dados não relacionados com a análise do risco de crédito do consumidor (art. 3º, § 3º, I), assim como dados pessoais sensíveis, entendidos como aqueles pertinentes "à origem social e étnica, à saúde, à informação genética, à orientação sexual e às convicções políticas, religiosas e filosóficas", nos termos do art. 3º, § 2º, II[297].

Em conjunto, o Código de Defesa do Consumidor e a Lei do Cadastro Positivo formam um robusto microssistema de proteção de dados pessoais aplicável às relações de consumo no Brasil. Ao longo das últimas três décadas, o poder judiciário tem crescentemente recorrido aos princípios e garantias previstos nesses diplomas para conformar limites às práticas de *credit scoring* e formação de bancos de dados sobre histórico de crédito.

Ao analisar a evolução da jurisprudência do Superior Tribunal de Justiça em

---

[296] BRASIL. *Lei Complementar nº 166, de 8 de abril de 2019*. Altera a Lei Complementar nº 105, de 10 de janeiro de 2001, e a Lei nº 12.414, de 9 de junho de 2011, para dispor sobre os cadastros positivos de crédito e regular a responsabilidade civil dos operadores. Brasília, DF: Presidência da República, 08 abr. 2019. Disponível em: http://www.planalto.gov.br/ccivil_03/LEIS/LCP/Lcp166.htm#art2. Acesso em: 8 abr. 2021.

[297] BRASIL. *Lei nº 12.414, de 09 de junho de 2011*. Disciplina a formação e consulta a bancos de dados com informações de adimplemento, de pessoas naturais ou de pessoas jurídicas, para formação do histórico de crédito. Brasília, DF: Presidência da República, 10 jun. 2011. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12414.htm. Acesso em: 17 ago. 2020.

matéria de proteção de dados pessoais, Ricardo Villas Bôas Cueva destaca que muito antes do advento de uma lei específica de proteção de dados pessoais no Brasil, já se discutia na jurisprudência do STJ a emergência de um novo conceito de privacidade, distinto daquele pautado na ideia de liberdade negativa.[298]

A formação dos primeiros precedentes nesta matéria teve início ainda na década de 1990, no contexto da recente entrada em vigor do Código de Defesa do Consumidor. Em 1995, o Superior Tribunal de Justiça julgou o RE nº 22.337-8/RS, Rel. Min. Ruy Rosado Aguiar, um dos primeiros precedentes no STJ sobre a formação de bancos de dados relativos a informações de crédito. O recorrido ajuizou inicialmente ação cautelar e, posteriormente, ação ordinária em face do Serviço de Proteção ao Crédito e da recorrente para cancelar seus registros negativos nos arquivos do SPC, tendo as ações sido julgadas procedentes e excluído o SPC do polo passivo. A recorrente, por sua vez, apelou das decisões, tendo o provimento ao apelo negado pela 7ª CC do Tribunal de Justiça do Rio Grande do Sul. Irresignada, a recorrente interpôs recurso especial, alegando, dentre outras razões, que a prescrição da ação de cobrança dos débitos que deram origem ao registro negativo só ocorreria em vinte anos, nos termos do art. 177 do Código Civil, e que os registros deveriam ser mantidos por igual período. O STJ decidiu, por unanimidade, não conhecer do recurso, ao reconhecer que o registro de dados pessoais no SPC deve ser cancelado após cinco anos, nos termos do art. 43, § 1º, do CDC.

O voto do Min. Rel. Ruy Rosado Aguiar é paradigmático em ao menos quatro aspectos. Primeiramente, ele constitui, como afirmado anteriormente, um dos primeiros precedentes no STJ sobre registros envolvendo informações de crédito. Em segundo lugar, seu voto torna clara a crescente preocupação da doutrina e jurisprudência brasileiras já naquele período com a proteção dos direitos fundamentais e dos direitos da personalidade no contexto do advento de novas tecnologias e do risco trazido pelas aumentadas capacidades de produção, coleta e tratamento de dados pessoais:

> A inserção de dados pessoais do cidadão em bancos de informações tem se constituído em uma das preocupações do Estado moderno, onde o uso da informática e a possibilidade de controle unificado das diversas atividades da pessoa, nas múltiplas situações de vida,

[298] CUEVA, R. V. B. A proteção dos dados pessoais na jurisprudência do Superior Tribunal de Justiça. *In:* TEPEDINO, G.; FRAZÃO, A.; OLIVA, M. D. *Lei Geral de Proteção de Dados Pessoais e suas repercussões no Direito Brasileiro*. 2. ed. São Paulo: Thomson Reuters, 2020. p. 86.

> permite o conhecimento de sua conduta pública e privada, até nos mínimos detalhes, podendo chegar à devassa de atos pessoais, invadindo área que deveria ficar restrita à sua intimidade; ao mesmo tempo, o cidadão objeto dessa indiscriminada colheita de informações, muitas vezes, sequer sabe da existência de tal atividade, ou não dispõe de eficazes meios para conhecer o seu resultado, retificá-lo ou cancelá-lo. E assim como o conjunto dessas informações pode ser usado para fins lícitos, públicos ou privados, na prevenção ou repressão de delitos, ou habilitando o particular a celebrar contratos com pleno conhecimento de causa, também pode servir, ao Estado e ao particular, para alcançar fins contrários à moral ou ao direito, como instrumento de perseguição política ou opressão econômica.[299]

Terceiro, chama a atenção o diálogo estabelecido com a doutrina e jurisprudência alemãs. Ao fundamentar o seu voto, o Min. Ruy Rosado faz menção expressa ao status de direito fundamental concedido à matéria no direito alemão, onde reconheceu-se ao cidadão o direito à "autodeterminação informacional":

> Nos países mais adiantados, algumas providências já foram adotadas. Na Alemanha, por exemplo, a questão está posta no nível das garantias fundamentais, com o direito de autodeterminação informacional (o cidadão tem o direito de saber quem sabe o que sobre ele), além da instituição de órgão independentes, à semelhança do *ombudsman*, com poderes para fiscalizar o registro de dados informatizados, pelos órgãos públicos e privados, para a garantia dos limites permitidos na legislação.[300]

Conforme assinala Cueva, esse novo conceito passou a aparecer em acórdãos relacionados à aplicação do art. 43 do CDC e aos cadastros negativos de crédito, tendo sido estas as primeiras oportunidades em que o STJ se manifestou sobre o tema.[301] Esse julgado, portanto, é ponto relevante no esforço de reconstrução da genealogia do conceito de autodeterminação informacional no Brasil, que, posteriormente, veio a consolidar-se não só na jurisprudência do STJ, como também na jurisprudência do Supremo Tribunal Federal, quando do julgamento da ADI nº 6.387/DF[302], e nos fundamentos da Lei Geral de Proteção de Dados, onde encontra-

---

[299] BRASIL. Superior Tribunal de Justiça. *Recurso Especial nº 22.337 – RS*. Relator: Ruy Rosado Aguiar. Data de Publicação: 20/03/1995. *In: R. Sup. Trib. Just. Brasília*, a.8 (77), jan. 1996. p. 206.

[300] BRASIL. Superior Tribunal de Justiça. *Recurso Especial nº 22.337 – RS*. Relator: Ruy Rosado Aguiar. Data de Publicação: 20/03/1995. *In: R. Sup. Trib. Just. Brasília*, a.8 (77), jan. 1996. p. 206.

[301] CUEVA, R. V. B. A proteção dos dados pessoais na jurisprudência do Superior Tribunal de Justiça. *In:* TEPEDINO, G.; FRAZÃO, A.; OLIVA, M. D. *Lei Geral de Proteção de Dados Pessoais e suas repercussões no Direito Brasileiro*. 2. ed. São Paulo: Thomson Reuters, 2020. p. 87.

[302] BRASIL. Supremo Tribunal Federal. *ADI nº 6.387/DF*. Rel. Min. Rosa Weber. Data de Publicação: 12/11/2020. Disponível em: http://portal.stf.jus.br/processos/detalhe.asp?incidente=5895165. Acesso em: 12 abr. 2021.

se positivado em seu art. 2º, II. Além de valer-se da experiência alemã como reforço argumentativo, no referido REsp, o Min. Ruy Rosado Aguiar associou o direito de acesso e de retificação de registros previsto no art. 43 do CDC à tutela da intimidade e da vida privada prevista no art. 5º, X e XII, da CF.

Por fim, é possível afirmar que esse precedente é paradigmático ao, de um lado, reconhecer a importância e relevância dos sistemas de proteção ao crédito para o bom funcionamento do mercado e como importante mecanismo no atendimento às legítimas expectativas dos credores, e, de outro, ao reconhecer a necessidade do estabelecimento de limitações e salvaguardas ao seu emprego:

> O Serviço de Proteção ao Crédito, instituído em diversas cidades pelas entidades de classe de comerciantes e lojistas, tem a finalidade de informar seus associados sobre a existência de débitos pendentes por comprador que pretenda obter novo financiamento. É evidente o benefício que dele decorre em favor da agilidade e da segurança das operações comerciais, assim como não se pode negar ao vendedor o direito de informar-se sobre o crédito do seu cliente na praça, e de repartir com os demais os dados que sobre ele dispõe. Essa atividade, porém, em razão da sua própria importância social e dos graves efeitos decorrentes deve ser exercido dentro dos limites que, permitindo a realização de sua finalidade, não se transforme em causa e ocasião de dano social maior do que o bem visado. É preciso admitir que tal registro somente deve ser feito com o prévio conhecimento do interessado, a fim de habilitá-lo a tomar as medidas cabíveis, fundadas na defesa que tiver, inclusive de inexistência do débito. Depois, impende considerar que tal registro não pode ser perpétuo. O nosso sistema jurídico não autoriza a indefinida permanência dos registros negativos nem para as sentenças criminais condenatórias, cujos efeitos desaparecem pelo simples efeito do tempo, daí a razão pela qual a Lei 8.078, de 11.9.90, no seu artigo 43, dispõe. Antes dele, a Súmula n. 11, do T JRS, dispunha: "A inscrição do nome do devedor do SPC pode ser cancelada após o decurso do prazo de três anos", a qual veio a ser alterado, estendendo o prazo para cinco anos.[303]

Desde então, o STJ passou a decidir sobre uma série de casos envolvendo a análise dos cadastros negativos e positivos de crédito e sobre a utilização de sistemas de avaliação de risco de crédito (*credit scoring*). Entre 2001 e 2010, Cueva chama a atenção para dois casos analisados pelo STJ: em 2001, no julgamento do REsp nº 306.570/SP, Rel. Min. Eliana Calmon, a corte reconheceu que "[...] o contribuinte ou o titular da conta bancária tem direito à privacidade em relação aos seus dados pessoais." Em 2010, ao analisar o REsp nº 1.168.547/RJ, Rel. Min. Luís Felipe

[303] BRASIL. Superior Tribunal de Justiça. *Recurso Especial nº 22.337 – RS*. Relator: Ruy Rosado Aguiar. Data de Publicação: 20/03/1995. *In: R. Sup. Trib. Just. Brasília*, a.8 (77), jan. 1996. p. 207.

Salomão, "[...] assentou-se a existência de um novo conceito de privacidade, bem como a necessidade de consentimento do interessado 'para dispor com exclusividade sobre as próprias informações, nelas incluindo o direito à imagem.'"[304]

Com o advento da Lei do Cadastro Positivo, em 2011, que ampliou o regime legal trazido pelo Código de Defesa do Consumidor, o STJ passou a examinar também questões relativas ao cadastro positivo de crédito. Neste contexto, destaca-se o REsp nº 1.348.532/SP, Rel. Min. Luís Felipe Salomão, no âmbito do qual reconheceu-se a abusividade de cláusula de adesão contida em contrato de prestação de serviço de cartão de crédito que previa o compartilhamento dos dados do titular com terceiros, sem que este pudesse se opor ao compartilhamento.[305]

Conforme assinala Cueva, com o advento da Lei do Cadastro Positivo, um dos grandes desafios enfrentados pelo STJ foram os julgamentos de inúmeros casos envolvendo sistemas de avaliação de risco de crédito (*credit-score*).[306] De acordo com Danilo Doneda, em 2016, havia cerca de 250 mil ações judiciais no Brasil sobre o assunto — sendo 80 mil delas apenas no estado do Rio Grande do Sul — nas quais os consumidores buscavam ser indenizados em razão dos sistemas de pontuação (e, em alguns casos, pela mera existência da pontuação).[307] Esse cenário levou o STJ ao julgamento de três recursos repetitivos[308] — REsp nº 1.419.697/RS[309], REsp nº 1.457.199/RS[310] e o REsp nº 1.304.736/RS[311] — e à edição de uma súmula — a

---

[304] CUEVA, R. V. B. A proteção dos dados pessoais na jurisprudência do Superior Tribunal de Justiça. *In:* TEPEDINO, G.; FRAZÃO, A.; OLIVA, M. D. *Lei Geral de Proteção de Dados Pessoais e suas repercussões no Direito Brasileiro*. 2. ed. São Paulo: Thomson Reuters, 2020. p. 87.

[305] Ibidem, p. 87-88.

[306] Ibidem, p. 88.

[307] DONEDA, Danilo. Current Judicial and Administrative Issues of Consumer Data Protection in Brazil. *In*: METZ, R.; BINDING, J.; HAIFEND, P. (Eds.). *Consumer Protection in Brazil, China and Germany:* a comparative study. Göttingen: Göttingen University Press, 2016. Disponível em: https://www.univerlag.uni-goettingen.de/bitstream/handle/3/isbn-978-3-86395-236-5/Huber_consumer.pdf?sequence=1&isAllowed=y. Acesso em: 08 abr. 2021. p. 154.

[308] "O caso foi decidido por um tipo especial de recurso, chamado 'recurso especial repetitivo'. No Código de Processo Civil, isso significa que a interpretação definida por essa decisão aplica-se a todos os casos semelhantes que compartilhem os mesmos elementos factuais. O julgamento ocorre por amostragem, o que significa que um caso é escolhido como 'representativo' de outros [...]." (ZANATTA, R. A. F. *Pontuação de Crédito e Direitos dos Consumidores*: o desafio brasileiro. São Paulo: Idec, 2017. p. 13).

[309] BRASIL. Superior Tribunal de Justiça. *Recurso Especial nº 1.419.697/RS*. Relator: Min. Paulo de Tarso Sanseverino. Julgado em: 12/11/2014. Data de Publicação: 17/11/2014. Disponível em: https://bdjur.stj.jus.br/jspui/bitstream/2011/114173/REsp1419697.pdf. Acesso em: 8 abr. 2021.

[310] BRASIL. *Recurso Especial nº 1.457.199/RS*. Relator: Min. Paulo de Tarso Sanseverino. Julgado em: 12/11/2014. Data de Publicação: 17/12/2014. Disponível em: https://bdjur.stj.jus.br/jspui/bitstream/2011/114173/REsp1457199.pdf. Acesso em: 8 abr. 2021.

[311] Idem. *Recurso Especial nº 1.304.736/RS*. Relator: Min. Luís Felipe Salomão. Julgado em: 24/02/2016. Data de Publicação: 20/03/2016. Disponível em:

Súmula 550 — entre o período de 2014 e 2016.

Em agosto de 2014, o STJ dava início à realização da primeira audiência pública de sua história[312], ocasião na qual se discutiu sobre o enquadramento jurídico da pontuação de crédito no ordenamento jurídico brasileiro: seria ela mera metodologia para análise de risco ou deveria ser caracterizada como espécie de banco de dados? O enquadramento se torna relevante em razão das consequências jurídicas que o acompanham: se enquadrado como banco de dados, seria necessário obter o consentimento do consumidor, conforme prevê a legislação consumerista. Essa era a interpretação sustentada, por exemplo, pelo Tribunal de Justiça do Rio Grande do Sul e pelo movimento brasileiro de defesa do consumidor. Já outros atores, como o Bacen e empresas do setor financeiro, sustentavam posição distinta: a pontuação de crédito seria apenas uma metodologia para cálculo do risco de crédito e a obtenção do consentimento específico para fazer esse cálculo não seria necessária.[313]

A decisão proferida no RE nº 1.419.167/RS foi no sentido de reconhecer que a pontuação de crédito é mera metodologia de análise de risco, o que, a partir de uma perspectiva garantista, foi encarado como uma derrota para a defesa dos consumidores no Brasil. Não obstante, a decisão cuidou também de estabelecer algumas salvaguardas legais: a pontuação de crédito só poderá ser considerada legal se observados o direito de transparência e a boa-fé, e a utilização de informações sensíveis, excessivas ou incorretas pode ensejar a indenização por danos morais ao consumidor. Nesse sentido, Zanatta argumenta que a decisão pode ser considerada equilibrada, uma vez que também cuidou de declarar um conjunto de limitações ao emprego de referidas metodologias e de direitos aos consumidores, como o direito de acesso aos dados e "novos direitos de transparência".[314]

Em conjunto, o REsp nº 1.419.697/RS e o REsp nº 1.457.199/RS sedimentaram um conjunto de 5 teses aplicáveis nos casos envolvendo pontuação de crédito:

> 1) O sistema "*credit scoring*" é um método desenvolvido para avaliação do risco de concessão de crédito, a partir de modelos estatísticos,

---

https://processo.stj.jus.br/processo/revista/documento/mediado/?componente=ITA&sequencial=1489563&num_registro=201200318393&data=20160330&peticao_numero=-1&formato=PDF. Acesso em: 8 abr. 2021.

[312] ZANATTA, R. A. F. *Pontuação de Crédito e Direitos dos Consumidores*: o desafio brasileiro. São Paulo: Idec, 2017.

[313] Ibidem, p. 3.

[314] Ibidem, p. 3-4.

> considerando diversas variáveis, com atribuição de uma pontuação ao consumidor avaliado (nota do risco de crédito).
> 2) Essa prática comercial é lícita, estando autorizada pelo art. 5º, IV, e pelo art. 7º, I, da Lei n. 12.414/2011 (lei do cadastro positivo).
> 3) Na avaliação do risco de crédito, devem ser respeitados os limites estabelecidos pelo sistema de proteção do consumidor no sentido da tutela da privacidade e da máxima transparência nas relações negociais, conforme previsão do CDC e da Lei n. 12.414/2011.
> 4) Apesar de desnecessário o consentimento do consumidor consultado, devem ser a ele fornecidos esclarecimentos, caso solicitados, acerca das fontes dos dados considerados (histórico de crédito), bem como as informações pessoais valoradas.
> 5) O desrespeito aos limites legais na utilização do sistema "*credit scoring*", configurando abuso no exercício desse direito (art. 187 do CC), pode ensejar a responsabilidade objetiva e solidária do fornecedor do serviço, do responsável pelo banco de dados, da fonte e do consulente (art. 16 da Lei n. 12.414/2011) pela ocorrência de danos morais nas hipóteses de utilização de informações excessivas ou sensíveis (art. 3º, § 3º, I e II, da Lei n. 12.414/2011), bem como nos casos de comprovada recusa indevida de crédito pelo uso de dados incorretos ou desatualizados.

Um apontamento relevante trazido por Zanatta quanto a esta decisão diz respeito ao seu caráter paradigmático e à possibilidade de extensão de seus efeitos para além dos limites dos sistemas de pontuação de crédito. Argumenta que, se a pontuação de crédito pode ser entendida como uma metodologia que emprega sistemas algorítmicos, é possível considerar a decisão proferida no âmbito do REsp nº 1.419.697/RS como a primeira decisão a tratar sobre *accountability* algorítmica e direitos dos consumidores no Brasil.[315] De acordo com o autor, "A decisão do STJ de 2014 não é simplesmente sobre os direitos dos consumidores. Trata de algoritmos e direitos à transparência".[316] Nesse sentido, ao estabelecer um novo direito coletivo de transparência, a decisão incide não apenas sobre a pontuação de crédito e sobre a proteção do consumidor, podendo também ser usada estrategicamente para debater a "*accountability* de algoritmos" no país.[317] Nas palavras do autor:

> A questão interessante para o futuro é avaliar se essa interpretação pode ser expandida e aplicada a outros casos que lidam com a relação entre consumidores e empresas privadas através da mediação de algoritmos. Pode-se argumentar que a decisão do STJ fornece fundamentos legais para identificar todo um conjunto de direitos à transparência para consumidores e decisões automatizadas. Isso

[315] ZANATTA, R. A. F. *Pontuação de Crédito e Direitos dos Consumidores*: o desafio brasileiro. São Paulo: Idec, 2017. p. 4.
[316] Ibidem, p. 22.
[317] ZANATTA, R. A. F. *Pontuação de Crédito e Direitos dos Consumidores*: o desafio brasileiro. São Paulo: Idec, 2017. p. 22.

> ocorre porque o Tribunal reconheceu que, apesar de não ser um novo "banco de dados" — como identificado na Lei do Cadastro Positivo — , consumidores possuem direitos à transparência quando lidam com algoritmos como em sistemas de pontuação de crédito, incluindo o direito de entender por quê uma decisão automatizada foi tomada. Veremos futuramente como essa decisão pode ser aplicada a outros algoritmos e decisões automatizadas.[318]
>
> O *tour de force* interpretativo fornecido pelo Ministro Paulo de Tarso Sanseverino deve ser claramente entendido por todos os ativistas de direitos digitais, porque fornece um sólido entendimento para futuros casos envolvendo coleta de dados, algoritmos e direitos à transparência.[319]

Como visto, o STJ reconheceu a legalidade do uso de dados pessoais, sem o consentimento do indivíduo, para fins de análise de risco de crédito e concluiu que essa prática é lícita e legítima, desde que presentes os fatores limitadores descritos acima e garantidos os direitos do consumidor, conformando um direito que se assemelha bastante a um possível direito à explicação, ao exigir a garantia da tutela da privacidade e máxima transparência nas relações contratuais, bem como ao conferir ao consumidor a possibilidade de requisitar esclarecimentos acerca das fontes dos dados considerados, bem como as informações pessoais valoradas para a análise de risco.[320] Em conjunto, essas decisões levaram à edição da Súmula 550, editada em 2015, que prescreve:

> Súmula 550. A utilização de escore de crédito, método estatístico de avaliação de risco que não constitui banco de dados, dispensa o consentimento do consumidor, que terá o direito de solicitar esclarecimentos sobre as informações pessoais valoradas e as fontes dos dados considerados no respectivo cálculo.[321]

Posteriormente, em 2016, no REsp nº 1.304.736/RS, a mesma corte julgou se o direito de acesso às fontes dos dados e a explicação da lógica do seu tratamento encontravam algum fator limitador[322]. Concluiu que existe interesse de agir do

---

[318] ZANATTA, R. A. F. *Pontuação de Crédito e Direitos dos Consumidores*: o desafio brasileiro. São Paulo: Idec, 2017. p. 16.
[319] Ibidem, p. 22.
[320] BRASIL. Superior Tribunal de Justiça. *Recurso Especial nº 1.419.697/RS*. Relator: Min. Paulo de Tarso Sanseverino. Julgado em: 12/11/2014. Data de Publicação: 17/11/2014. Disponível em: https://bdjur.stj.jus.br/jspui/bitstream/2011/114173/REsp1419697.pdf. Acesso em: 8 abr. 2021.
[321] Idem. *Súmula 550*. Julgado em: 14/10/2015. Data de Publicação: 19/10/2015.
[322] BRASIL. *Recurso Especial nº 1.304.736/RS*. Relator: Min. Luís Felipe Salomão. Julgado em: 24/02/2016. Data de Publicação: 20/03/2016. Disponível em: https://processo.stj.jus.br/processo/revista/documento/mediado/?componente=ITA&sequencial=14895

consumidor que deseja conhecer os principais elementos e critérios considerados para a análise do seu histórico e as informações pessoais utilizadas – respeitado o segredo empresarial – desde que tenha sido atingido por tais critérios quando tentou obter crédito no mercado,[323] p. ex., deixou de conseguir crédito devido à pontuação que lhe foi atribuída. Na ocasião, fixou-se a seguinte tese:

> Em relação ao sistema *credit scoring*, o interesse de agir para a propositura da ação cautelar de exibição de documentos exige, no mínimo, a prova de: i) requerimento para obtenção dos dados ou, ao menos, a tentativa de fazê-lo à instituição responsável pelo sistema de pontuação, com a fixação de prazo razoável para atendimento; e ii) que a recusa do crédito almejado ocorreu em razão da pontuação que lhe foi atribuída pelo sistema Scoring.[324]

O STJ estabeleceu, assim, um critério que até então não encontrava respaldo na lei, possibilitando reconhecer a existência do direito à explicação de decisões totalmente automatizadas, desde que tais decisões tenham um impacto específico na vida das pessoas.

Dessa forma, em conjunto com o CDC, a Lei do Cadastro Positivo forma um microssistema de proteção de dados pessoais que, infelizmente, restringe-se apenas ao caso da concessão de crédito, embora haja uma construção interpretativa no sentido de reconhecer a possibilidade de extensão dessas salvaguardas para além das situações envolvendo análise e pontuação de crédito, ampliando-a também para situações envolvendo o emprego de sistemas automatizados no Brasil, conforme apontado anteriormente. Nessas situações, o consumidor pode requisitar informações sobre o uso de seus dados em uma decisão automatizada de classificação de risco para concessão ou não do crédito. Caso não concorde com esta decisão por entender que foi tomada em desacordo com os critérios estabelecidos na legislação e na jurisprudência, pode pedir a sua revisão por uma pessoa, conforme garantido no rol de direitos listado acima. A revisão humana, em tese, afastaria os elementos que foram indevidamente utilizados pelo algoritmo, como dados em excesso ou dados

63&num_registro=201200318393&data=20160330&peticao_numero=-1&formato=PDF. Acesso em: 8 abr. 2021.

[323] MENDES, L. S. *Privacidade, proteção de dados e defesa do consumidor*: linhas gerais de um novo direito fundamental. São Paulo: Saraiva, 2014.

[324] BRASIL. Superior Tribunal de Justiça. *Recurso Especial nº 1.304.736/RS*. Relator: Min. Luís Felipe Salomão. Julgado em: 24/02/2016. Data de Publicação: 20/03/2016. Disponível em: https://processo.stj.jus.br/processo/revista/documento/mediado/?componente=ITA&sequencial=1489563&num_registro=201200318393&data=20160330&peticao_numero=-1&formato=PDF. Acesso em: 8 abr. 2021.

sensíveis. Todavia, cabe questionar se uma decisão tomada por uma pessoa, e não por um sistema, seria menos enviesada.[325]

Nota-se, ainda, que os direitos e balizas previstos nas leis e precedentes judiciais foram absorvidos pela Lei Geral de Proteção de Dados do Brasil, o que sugere que o racional detalhado e analisado pelo Superior Tribunal de Justiça também deve ser utilizado para interpretar a LGPD, a qual analisaremos no tópico a seguir.

3.2.2 A regulamentação específica da LGPD

Como visto, o Código de Defesa do Consumidor e a Lei do Cadastro Positivo regulamentam o direito à explicação e à revisão de decisões automatizadas no âmbito das relações de consumo, mais especificamente quando envolvem a concessão de crédito e cálculo de risco de inadimplência. Mas essa proteção setorial é insuficiente. Na verdade, em nenhum dos exemplos mencionados neste trabalho os instrumentos de proteção consumerista seriam satisfatórios. Daí a importância de previsões capazes de expandir esses direitos para contextos mais variados, *online* e *off-line*, envolvendo o uso de dados pessoais. Para endereçar essa necessidade, foi promulgada a Lei Geral de Proteção de Dados do Brasil, a LGPD, que estrutura o regime da proteção de dados pessoais no ordenamento brasileiro em torno de um conjunto normativo unitário[326] e transversal[327], transplantando o sistema setorial de proteção nacional para um geral, que abrange o tratamento de dados pessoais, independente do contexto, setor e mercado.

A Lei nº 13.709/2018 complementa, harmoniza e unifica um ecossistema de mais de quarenta normas setoriais que regulam, de forma direta e indireta, a proteção da privacidade e dos dados pessoais no Brasil[328]. Foi inspirada nas discussões que culminaram no regulamento europeu de proteção de dados e tem por objetivo não só

---

[325] MILLER, J. M. *Dignity as a New Framework, Replacing the Right to Privacy*. Rochester, NY: Social Science Research Network, 2008. Disponível em: https://papers.ssrn.com/abstract=1127986. Acesso em: 27 maio. 2020.

[326] DONEDA, D. *Da privacidade à proteção de dados pessoais:* fundamentos da Lei Geral de Proteção de dados. 2. ed. São Paulo: Thomson Reuters Brasil, 2019. p. 259.

[327] WIMMER, M. Os desafios do *enforcement* na LGPD: fiscalização, aplicação de sanções administrativas e coordenação intergovernamental. *In:* MENDES, L. S.; DONEDA, D.; SARLET, I. W.; RODRIGUES JR., O. L. (Coords.). *Tratado de Proteção de Dados Pessoais*. Rio de Janeiro: Forense, 2021. p. 376-377.

[328] MONTEIRO, R. L. Lei Geral de Proteção de Dados do Brasil: uma análise detalhada. *Jota*, 14 jul. 2018. Disponível em: https://www.jota.info/opiniao-e-analise/colunas/agenda-da-privacidade-e-da-protecao-de-dados/lgpd-analise-detalhada-14072018. Acesso em: 07 abr. 2021.

conferir às pessoas maior controle sobre seus dados, mas também fomentar um ambiente de desenvolvimento econômico e tecnológico, mediante regras flexíveis e adequadas para lidar com os mais inovadores modelos de negócio baseados no uso de dados pessoais. Isso inclui modelos de negócio que se valem de algoritmos para auxiliar na tomada de decisões automatizadas.

A LGPD tem por objetivo, dentre outros, equilibrar interesses econômicos e sociais, garantindo a continuidade desse tipo de processo decisório, mas também limitando eventuais abusos, fazendo-o por meio da instituição de mecanismos aptos a diminuir a assimetria informacional, e, por consequência, de poder, entre os titulares sujeitos a decisões automatizadas e os agentes de tratamento responsáveis pelos sistemas automatizados.

Nos subtópicos a seguir, apresentaremos a regulamentação específica da LGPD em matéria de decisões automatizadas em três momentos. Primeiramente, destacaremos as disposições da lei relevantes no contexto de decisões automatizadas, elencando o conjunto de direitos e princípios consubstanciados no texto que podem ser tomados como elementos estruturantes de um direito à explicação. Em seguida, apresentaremos o debate brasileiro em torno de decisões automatizadas e sobre a possibilidade de reconhecimento da existência de um direito à explicação na LGPD, a partir dos principais e mais recentes trabalhos publicados sobre o tema. Por fim, colocaremos em perspectiva os regimes da LGPD e da GDPR, buscando analisar em que medida e em quais aspectos o direito à explicação que pode ser derivado de uma interpretação sistemática da LGPD seria mais abrangente ou mais restrito do que aquele previsto na normativa europeia.

3.2.2.1 O regime jurídico da LGPD aplicável às decisões automatizadas

A Lei nº 13.709/2018, justamente em razão de seu caráter geral e transversal, adotou inúmeras disposições de caráter normativo aberto, de conteúdo essencialmente principiológico. Neste sentido, a lei estabelece objetivos, fundamentos e princípios que deverão ser observados em qualquer atividade de tratamento de dados. Algumas destas disposições possuem especial relevância no contexto de decisões automatizadas baseadas no tratamento de dados pessoais.

Em seu art. 1º, a lei elenca como objetivo a proteção dos direitos fundamentais de liberdade e de privacidade e o livre desenvolvimento da personalidade da pessoa

natural. No contexto de crescente ubiquidade de sistemas automatizados que operam de forma cada vez mais invasiva, atuando na distribuição e alocação de bens jurídicos e impactando, positiva ou negativamente, no livre desenvolvimento da personalidade dos indivíduos[329], faz-se necessário ter este objetivo no horizonte quando tratamos da regulação de decisões automatizadas. Já em seu art. 2º, a lei dispõe que a disciplina da proteção de dados pessoais tem como fundamentos, dentre outros, o respeito à privacidade (art. 2º, I), a autodeterminação informativa (art. 2º, II), a liberdade de expressão, de informação, de comunicação e de opinião (art. 2, III), o desenvolvimento tecnológico e a inovação (art. 2º, V), a livre iniciativa, a livre concorrência e a defesa do consumidor (art. 2º, VI) e os direitos humanos, o livre desenvolvimento da personalidade, a dignidade e o exercício da cidadania pelas pessoas naturais (art. 2º, VII). Juntos, os objetivos e fundamentos da Lei Geral de Proteção de Dados devem ser tomados como vetores interpretativos e conformadores das atividades de tratamento de dados pessoais, sobretudo naquelas em que o risco de comprometimento destes objetivos e fundamentos é expressivo, como é o caso de sistemas automatizados baseados em dados.

Em seu art. 6º, além da cláusula geral de boa-fé, a lei prevê ainda um robusto conjunto de princípios que deverão ser observados em todas as atividades de tratamento de dados pessoais dentro do escopo da LGPD, quais sejam: princípio da finalidade (art. 6º, I), princípio da adequação (art. 6º, II), princípio da necessidade (art. 6º, III), princípio do livre acesso (art. 6º, IV), princípio da qualidade dos dados (art. 6º, V), princípio da transparência (art. 6º, VI), princípio da segurança (art. 6º, VII), princípio da prevenção (art. 6º, VIII), princípio da não discriminação (art. 6º, IX) e princípio da responsabilização e prestação de contas (art. 6º, X).

No contexto de decisões automatizadas, que, por diferentes razões, sejam elas técnicas, por opção de *design* ou por limitações legais, encontram-se crescentemente cercados de opacidade, cabe destacar, primeiramente, o princípio da transparência. O art. 6º, VI, da LGPD, confere aos titulares a garantia de obtenção de informações claras, precisas e facilmente acessíveis sobre a realização do tratamento e os respectivos agentes de tratamento[330], observados os segredos comercial e industrial.

---

[329] BIONI, B. R. *Proteção de dados pessoais*: a função e os limites do consentimento. 2. ed. Rio de Janeiro: Forense, 2021. p. 55-59.

[330] "Art. 5º. Para os fins desta Lei, considera-se: [...] controlador: pessoa natural ou jurídica, de direito público ou privado, a quem competem as decisões referentes ao tratamento de dados pessoais; [...]

Ou seja, a garantia para que se requisite de órgãos públicos e privados informações sobre como os seus dados são usados. Esse princípio, que dá origem ao direito de acesso aos dados, é complementado pelo art. 19, que dispõe que “A confirmação de existência ou o acesso a dados pessoais serão providenciados, mediante requisição do titular" e se dará ou por forma simplificada, imediatamente (art. 19, I), ou

> [...] por meio de declaração clara e completa, que indique a origem dos dados, a inexistência de registro, os critérios utilizados e a finalidade do tratamento, observados os segredos comercial e industrial, fornecida no prazo de até 15 (quinze) dias, contado da data do requerimento do titular.[331]

Ou seja, o princípio da transparência deve reger toda e qualquer relação do responsável pelo tratamento de dados pessoais com o titular dos dados, garantindo a este o direito de acesso aos seus dados pessoais. Esse princípio também pressupõe o dever de informar o titular sobre os critérios e procedimentos utilizados para se chegar à determinada decisão automatizada (art. 20, § 2º), podendo ser lido como um dos elementos de sustentação para o reconhecimento de um direito à explicação na LGPD [332].

De igual modo, o princípio do livre acesso, ao garantir aos titulares a consulta facilitada e gratuita sobre a forma e a duração do tratamento, bem como sobre a integralidade de seus dados pessoais, também é de grande relevância no contexto de decisões automatizadas e soma-se como mais um elemento apto a sustentar o reconhecimento de um direito à explicação na LGPD.

O art. 6º, V, enuncia o princípio da qualidade dos dados, por meio do qual garante-se aos titulares a exatidão, clareza, relevância e atualização dos dados, de acordo com a necessidade e para o cumprimento da finalidade de seu tratamento. Como vimos no Capítulo 1, decisões automatizadas são construídas a partir dos dados de *input*, e a qualidade dos *outputs*, ou seja, das decisões resultantes, está

---

operador: pessoa natural ou jurídica, de direito público ou privado, que realiza o tratamento de dados pessoais em nome do controlador; [...] agentes de tratamento: o controlador e o operador.” (BRASIL. *Lei nº 13.709, de 14 de agosto de 2018*. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF: Presidência da República, 14 ago. 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015- 2018/2018/lei/L13709.htm. Acesso em: 07. nov. 2021).

[331] BRASIL. *Lei nº 13.709, de 14 de agosto de 2018*. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF: Presidência da República, 14 ago. 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015- 2018/2018/lei/L13709.htm. Acesso em: 07. nov. 2020.

[332] Há uma clara convergência entre a redação dos arts. 6º, VI, que enuncia o princípio da transparência, a redação do art. 20 da lei, que trata sobre decisões automatizadas, e àquela do inciso II do art. 19, que trata sobre o direito de acesso.

diretamente relacionada à qualidade dos dados empregados. Neste sentido, a não observância do princípio da qualidade pode dar origem ou reforçar vieses algorítmicos discriminatórios, impactando a esfera de direitos do titular ou mesmo de grupos de sujeitos não diretamente identificados. Não discriminação, inclusive, constitui outro princípio a ser estritamente observado no contexto de decisões automatizadas, uma vez que a LGPD veda a realização de tratamento com fins discriminatórios ilícitos ou abusivos (art. 6º, IX).

Por fim, cabe reconhecer ainda a pertinência dos princípios da prevenção (art. 6º, VIII), que preconiza a adoção de medidas para prevenir a ocorrência de danos em virtude do tratamento de dados pessoais, e da responsabilização e prestação de contas (art. 6º, X), que exige a demonstração, pelo agente, da adoção de medidas eficazes e capazes de comprovar a observância e o cumprimento das normas de proteção de dados pessoais e, inclusive, da eficácia dessas medidas, ambos de extrema relevância no contexto de atividades altamente cercadas de risco como o são as decisões automatizadas.

No tocante ao plexo de direitos positivados, cabe destacar os direitos de confirmação da existência de tratamento e de acesso aos dados (art. 18, I e II, art. 9º e art. 19), já mencionados, e o art. 20 da LGPD, que garante o direito de solicitar a revisão de uma decisão tomada unicamente com base em tratamento automatizado. O objetivo aqui é evitar que indivíduos sejam alvo de práticas discriminatórias dos algoritmos responsáveis pela decisão.

Além destes artigos, cabe destacar o tratamento dado pela lei ao caso de reidentificação de dados anonimizados[333]. Quando utilizados para composição de um perfil comportamental, dados dessa natureza poderão ser considerados como pessoais, desde que façam referência a uma pessoa identificada. De acordo com o art. 12, dados anonimizados somente são considerados dados pessoais "[...] quando o processo de anonimização ao qual foram submetidos for revertido, utilizando exclusivamente meios próprios, ou quando, com esforços razoáveis, puder ser

[333] "Art. 5º. Para os fins desta Lei, considera-se:
III — dado anonimizado: dado relativo a titular que não possa ser identificado, considerando a utilização de meios técnicos razoáveis e disponíveis na ocasião de seu tratamento;
XI — anonimização: utilização de meios técnicos razoáveis e disponíveis no momento do tratamento, por meio dos quais um dado perde a possibilidade de associação, direta ou indireta, a um indivíduo;" (BRASIL. *Lei nº 13.709, de 14 de agosto de 2018*. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF: Presidência da República, 14 ago. 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015- 2018/2018/lei/L13709.htm. Acesso em: 07. nov. 2021).

revertido", ou então quando se tratar de dados "utilizados para formação do perfil comportamental de determinada pessoa natural, se identificada" (art. 12, *caput* e § 2º).[334]

O que o artigo trata como "pessoa identificada" faz referência ao conceito de dado pessoal previsto no art. 5º, I, que pode incluir formas de diretamente identificar uma pessoa natural, por meio do seu nome ou características distintivas únicas; identificadores únicos, como CPF, RG, CNH; e até mesmo identificadores eletrônicos, como *e-mail* e *cookies*.

Ainda, nos termos do art. 20, § 1º, o controlador deverá fornecer, sempre que solicitadas, informações claras e adequadas a respeito dos critérios e dos procedimentos utilizados para a decisão automatizada, observados os segredos comerciais e industriais. Caso o responsável pelo processamento dos dados se recuse a fornecer os dados pessoais utilizados na decisão automatizada e a explicar os critérios e/ou a lógica subjacente dos algoritmos que controlam o processo de tomada de decisão, a Autoridade Nacional de Proteção de Dados ("ANPD") *poderá* realizar auditoria para verificação de aspectos discriminatórios em tratamento automatizado de dados pessoais, após processo administrativo em que deve ser garantido a ampla defesa e o contraditório[335], nos termos do art. 20, § 2º. Esse procedimento visa verificar, principalmente, a existência de aspectos discriminatórios, tais como o uso de dados pessoais sensíveis ou que excedam a finalidade pretendida[336]. Todavia, aferir eventuais discriminações pode ser um trabalho extremamente técnico, devido à complexidade dos algoritmos, o que demonstra a necessidade de a ANPD ter um corpo de profissionais altamente especializado e preparado. Portanto, a LGPD amplia essas vedações no uso de dados para além das

---

[334] BRASIL. *Lei nº 13.709, de 14 de agosto de 2018*. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF: Presidência da República, 14 ago. 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015- 2018/2018/lei/L13709.htm. Acesso em: 07. nov. 2020.

[335] "Art. 23 [...]. § 3º Os prazos e procedimentos para exercício dos direitos do titular perante o Poder Público observarão o disposto em legislação específica, em especial as disposições constantes da Lei nº 9.507, de 12 de novembro de 1997 (Lei do Habeas Data), da Lei nº 9.784, de 29 de janeiro de 1999 (Lei Geral do Processo Administrativo), e da Lei nº 12.527, de 18 de novembro de 2011 (Lei de Acesso à Informação)." (BRASIL. *Lei nº 13.709, de 14 de agosto de 2018*. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF: Presidência da República, 14 ago. 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015- 2018/2018/lei/L13709.htm. Acesso em: 07. nov. 2020).

[336] "Art. 20 [...]. § 2º Em caso de não oferecimento de informações de que trata o § 1º deste artigo baseado na observância de segredo comercial e industrial, a autoridade nacional poderá realizar auditoria para verificação de aspectos discriminatórios em tratamento automatizado de dados pessoais." (BRASIL. *Lei nº 13.709, de 14 de agosto de 2018*. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF: Presidência da República, 14 ago. 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015- 2018/2018/lei/L13709.htm. Acesso em: 07. nov. 2020).

relações de consumo, para incluir outros usos de dados pessoais.

Em síntese, a LGPD sistematiza um conjunto de objetivos, fundamentos, princípios e direitos que oferecem um robusto ferramental de proteção aos direitos do titular no contexto de decisões automatizadas, que podem ser entendidos como elementos aptos a sustentar a existência de um efetivo direito à explicação. Como vimos, a LGPD garante aos indivíduos o direito a ter acesso a informações sobre que tipos de dados pessoais são utilizados para alimentar algoritmos responsáveis por decisões automatizadas. Caso o processo automatizado tenha por finalidade formar perfis comportamentais ou se valha de um perfil comportamental para tomar uma decisão subsequente, essa previsão também incluirá o acesso aos dados anonimizados utilizados para enriquecer tais perfis. Esse direito ainda inclui a possibilidade de conhecer os critérios utilizados para tomar a decisão automatizada[337] e de solicitar a revisão da decisão quando esta afetar os interesses dos titulares.

Pela lei, os direitos à explicação e à revisão de decisões automatizadas podem ser usufruídos em qualquer tipo de tratamento de dados pessoais, independente do setor ou mercado, público ou privado. Isto confere ao titular dos dados pessoais ferramentas importantes para coibir abusos e práticas discriminatórias no uso dos seus dados. Tais direitos devem contribuir diretamente para uma mudança na forma como produtos, serviços e processos são desenvolvidos, devido às obrigações de informar e explicar atribuídas aos agentes de tratamento. Esses terão que pensar, desde a concepção, como garantir os direitos previstos na LGPD, o que deve – pelo menos assim se espera – diminuir a obscuridade e a opacidade dos algoritmos[338].

3.2.2.2 Decisões automatizadas e direito à explicação na LGPD: o debate brasileiro

A produção bibliográfica mais recente no cenário brasileiro sobre discriminação algorítmica, decisões automatizadas e o direito à explicação aponta para uma significativa convergência em torno de, ao menos, três aspectos: (i) assim como no cenário europeu, há incertezas sobre a efetiva existência de um direito à explicação, ao menos na amplitude semântica do termo, na Lei Geral de Proteção de Dados

[337] A Lei prevê duas exceções: os casos de segredo industrial e comercial. Observa-se que nessas hipóteses, é importante analisar caso a caso, uma vez que a lei não especifica critérios para determinar quando se trata do caso de segredo comercial/industrial.

[338] PASQUALE, F. *The Black box society:* the secret algorithms that control money and information. First Harvard University Press paperback edition ed. Cambridge, Massachusetts; London, England: Harvard University Press, 2015.

brasileira; (ii) ainda que a LGPD não preveja expressamente o direito à explicação, uma leitura sistemática da lei não impede que outros direitos sejam derivados da própria racionalidade adotada, bem como dos princípios, dos objetivos, dos fundamentos e dos outros direitos positivados; e (iii) o debate sobre o direito à explicação no Brasil ocorre em paralelo ao debate travado na UE, sendo largamente dele dependente tanto em termos de fundamentos teóricos quanto em termos de construções hermenêuticas; neste ponto, é possível observar que há uma inevitável comparação das disposições da LGPD e da GDPR, havendo o cuidado de, ao realizar essa comparação e transposição do debate, destacar as especificidades e os desafios próprios da lei brasileira.

Topograficamente, os direitos dos titulares encontram-se previstos no Capítulo III da LGPD, mais especificamente nos arts. 18 e 20, que conferem ao titular os direitos, que assim podem ser aglutinados, de confirmação da existência de tratamento e de acesso aos dados (art. 18, I e II), direito de retificação (art. 18, III), direito de cancelamento (art. 18, IV, VI e IX), direito de oposição (art. 18, § 2º), direito à explicação (interpretação sistemática), direito à revisão de decisões automatizadas (art. 20) e direito de portabilidade (art. 18, V).

Não obstante, parece-nos adequado reconhecer que, ainda que os direitos dos titulares se encontrem concentrados no Capítulo III da LGPD e, em especial, em seus arts. 18 e 20, é muito mais adequado concebê-los a partir de uma leitura sistemática e abrangente da normativa.[339] Nesse sentido, Frazão bem pontua que "[...] antes mesmo de adentrar no capítulo específico sobre os direitos dos titulares, a LGPD já havia traçado um robusto conjunto de direitos e garantias, que precisa ser considerado para se entender a real extensão do Capítulo III."[340]

Quanto ao referido, a autora afirma, ao discorrer sobre o art. 17, que assegura à pessoa natural a titularidade de seus dados e lhe garante os direitos fundamentais de liberdade, intimidade e privacidade, que, justamente por repetir alguns dos direitos já mencionados anteriormente no texto e não dispor expressamente em relação a outros, deverá ele ser interpretado em conformidade com os artigos que o antecedem,

---

[339] MONTEIRO, R. L.; CRUZ, S. N. e. Direitos dos titulares na LGPD: fundamentos, limites e aspectos práticos. *In*: FRANCOSKI, D. de S. L.; TASSO, F. A. (Coords.). *A Lei Geral de Proteção de Dados Pessoais*: Aspectos Práticos e Teóricos Relevantes no Setor Público e Privado. São Paulo: Thomsom Reuters, abr. 2021.

[340] FRAZÃO, A. Nova LGPD: direitos dos titulares de dados pessoais. *Jota*, 24 out. 2018. Disponível em: https://www.jota.info/opiniao-e-analise/colunas/constituicao-empresa-e-mercado/nova-lgpd-direitos-dos-titulares-de-dados-pessoais-24102018. Acesso em: 06 jan. 2021.

em especial os relativos "[...] ao livre desenvolvimento da personalidade, à autodeterminação informativa, à dignidade da pessoa humana e ao exercício da cidadania."[341] No mesmo sentido aponta Silva, para quem uma leitura sistemática da lei não impede que outros direitos sejam derivados dos princípios e da lógica adotada pela LGPD[342], sendo este o caso, por exemplo, do direito à explicação no ordenamento brasileiro.

Nesse aspecto, há uma clara convergência entre os autores acerca de quais seriam os elementos normativos da LGPD aptos a sustentar o reconhecimento de um direito à explicação. Seriam eles: os objetivos, fundamentos e princípios da lei, em especial a autodeterminação informativa (art. 2º, II) e os princípios de livre acesso, qualidade, transparência, não-discriminação e responsabilidade e prestação de contas (art. 6º, IV, V, VI, IX e X); o art. 19, que dispõe sobre os direitos de confirmação da existência de tratamento e de acesso; e o art. 20, que prevê os direitos de (i) obtenção de revisão da decisão da automatizada, (ii) de acesso a informações claras e adequadas acerca dos critério e procedimentos envolvidos em dada decisão automatizada e de (iii) petição junto à Autoridade Nacional de Proteção de Dados para realização de auditoria (art. 20, *caput* e § 1º).

Ana Frazão afirma que o art. 20 da LGPD consagra no ordenamento jurídico brasileiro um conjunto de direitos aos titulares de dados pessoais no contexto de decisões automatizadas, a saber: o direito de acesso e informação sobre os critérios e procedimentos que orientaram a decisão automatizada, o direito de oposição à decisão automatizada e de manifestar seu ponto de vista, o direito de obter uma revisão da decisão, e o direito de requisitar a realização de auditoria junto à Autoridade Nacional de Proteção de Dados.[343] Para a autora:

> [...] tais direitos decorrem não apenas da autodeterminação informacional do cidadão e do controle que a lei lhe atribui sobre os seus dados pessoais, como também de importantes princípios da LGPD, dentre os quais (i) o livre acesso (art. 6º, VI), a qualidade e a clareza dos dados (art. 6º, V), a transparência dos dados, o que requer

[341] FRAZÃO, A. Nova LGPD: direitos dos titulares de dados pessoais. *Jota*, 24 out. 2018. Disponível em: https://www.jota.info/opiniao-e-analise/colunas/constituicao-empresa-e-mercado/nova-lgpd-direitos-dos-titulares-de-dados-pessoais-24102018. Acesso em: 06 jan. 2021.
[342] SILVA, P. R. Os direitos dos titulares de dados. *In*: MULHOLLAND, Caitlin (Org.). *A LGPD e o novo marco normativo no Brasil*. Porto Alegre: Arquipélago, 2020. p. 195-196.
[343] FRAZÃO, A. O direito à explicação e à oposição diante de decisões totalmente automatizadas. *Jota*, 05 dez. 2018. Disponível em: https://www.jota.info/paywall?redirect_to=//www.jota.info/opiniao-e-analise/colunas/constituicaoempresa-e-mercado/o-direito-a-explicacao-e-a-oposicao-diante-de-decisoes-totalmente-automatizadas-05122018. Acesso em: 11 abr.2021.

> ‘informações claras, precisas e facilmente acessíveis (art. 6º, VI), (iv) a prevenção de danos (art. 6º, VII), já que, como se verá adiante, o tratamento totalmente automatizado envolve sérios riscos para os titulares, (v) a não discriminação (art. 6º, IX) e (vi) a responsabilização e prestação de contas (art. 6º, X). [344]

Esse entendimento é também compartilhado por outros(as) autores(as), como Mullholand e Frahjof, para quem esse conjunto de direitos decorre diretamente não apenas da autodeterminação informacional reconhecida aos indivíduos, “[...] mas também do atendimento aos princípios reconhecidos na LGPD, como o livre acesso (art. 6º, IV), a transparência dos dados (art. 6º, VI) e a não discriminação (art. 6º, IX).”[345] Na mesma direção aponta Silva, para quem os direitos à explicação e à revisão estariam relacionados ao princípio da transparência, e teriam como objetivo oferecer mecanismos com vistas a minimizar o impacto do crescente uso de algoritmos para realização de julgamentos e avaliações.[346]

Semelhante entendimento é apresentado por Souza, Perrone e Magrani, que argumentam ser possível fundar um direito à explicação na lei brasileira a partir de três elementos: o princípio da transparência, o direito de acesso à informação e a partir do entendimento de que o direito à explicação seria um pressuposto para o exercício de outros direitos.[347] Trataremos dos dois últimos elementos mais adiante. Por hora, cabe destacar que, para os autores, o direito à explicação derivaria, primeiramente, do princípio da transparência, que possui aplicação transversal, atuando como um eixo norteador de toda a LGPD. Reconhecem que, assim como no cenário europeu, o princípio da transparência permite sustentar e garantir ao titular o direito à explicação, fundamentando ainda um dever de transparência ativa por parte do controlador, independentemente de provocação do titular. Nesse sentido, afirmam:

> [...] o princípio da transparência, entendido na sua dimensão de

---

[344] FRAZÃO, A. O direito à explicação e à oposição diante de decisões totalmente automatizadas. *Jota*, 05 dez. 2018. Disponível em: https://www.jota.info/paywall?redirect_to=//www.jota.info/opiniao-e-analise/colunas/constituicaoempresa-e-mercado/o-direito-a-explicacao-e-a-oposicao-diante-de-decisoes-totalmente-automatizadas-05122018. Acesso em: 11 abr.2021.

[345] MULHOLLAND, C.; FRAJHOF, I. Z. Inteligência Artificial e a Lei Geral 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. (Coords.). *Inteligência Artificial e Direito:* Ética, Regulação e Responsabilidade. 2. ed. São Paulo: Revista dos Tribunais, 2020.

[346] SILVA, P. R. Os direitos dos titulares de dados. *In*: MULHOLLAND, Caitlin (Org.). *A LGPD e o novo marco normativo no Brasil*. Porto Alegre: Arquipélago, 2020. p. 210.

[347] SOUZA, C. A.; PERRONE, C.; MAGRANI, E. O direito à explicação entre a experiência europeia e a sua positivação na LGPD. *In*: MENDES, L. S.; DONEDA, D.; SARLET, I. W.; RODRIGUES JÚNIOR, O. L.; BIONI, B. (Orgs.). *Tratado de Proteção de Dados Pessoais*. Rio de Janeiro: Forense, 2021. p. 243-270. p. 261.

> explicação, permite o equilíbrio dos interesses econômicos e sociais. Por um lado, admite o uso de decisões automatizadas e, por outro, diminui a assimetria de informação entre os agentes públicos e privados e os indivíduos ao tornar obrigatória a prestação de informações para o titular.[348]

Ainda no rol de princípios fundantes do direito à explicação na LGPD, tem-se o princípio da não-discriminação, de relevância central no contexto de decisões automatizadas. Mattiuzo traz uma importante contribuição na tarefa de interpretação deste princípio na lei brasileira.

De acordo com a autora, apesar de o conceito de discriminação estar frequentemente associado a práticas de exclusão e segmentação, o seu correto entendimento deve partir de uma noção mais ampla de discriminação enquanto generalização, ou "ideia-conceito de discriminação", tendo em vista que "toda discriminação é, de alguma forma, generalizante."[349] Neste sentido, argumenta que o recurso a generalizações é bastante comum no nosso cotidiano, inclusive no campo do direito, sobretudo quando se faz necessário tornar factível a tomada de decisões em um contexto de informações altamente assimétricas. Em seguida, traz exemplos de como as generalizações foram incorporadas como ferramentas no ordenamento jurídico, como se dá, por exemplo, na definição de categorias jurídicas. Destaca, ainda, que as categorias jurídicas, justamente por serem generalizantes, possuem exceções e limitações, e que, havendo limitações na própria forma de classificação dos indivíduos, é perfeitamente possível que determinado indivíduo "[...] não seja corretamente descrito pela característica do grupo ao qual pertence", sendo neste sentido, que se pode afirmar que toda generalização é, de alguma maneira, uma forma de discriminação.[350]

Para a autora, contudo, nem toda generalização pode ser enquadrada como uma "conduta discriminatória", no sentido negativo do termo. Isso porque há dois tipos

---

[348] SOUZA, C. A.; PERRONE, C.; MAGRANI, E. O direito à explicação entre a experiência europeia e a sua positivação na LGPD. *In*: MENDES, L. S.; DONEDA, D.; SARLET, I. W.; RODRIGUES JÚNIOR, O. L.; BIONI, B. (Orgs.). *Tratado de Proteção de Dados Pessoais*. Rio de Janeiro: Forense, 2021. p. 243-270. p. 262.

[349] MATTIUZO, M. Discriminação algorítmica: reflexões no contexto da Lei Geral de Proteção de Dados Pessoais. *In*: DONEDA, D.; MENDES, L. S.; CUEVA, R. V. B. (Coords.). *Lei Geral de Proteção de Dados (Lei nº 13.709/2018) — A caminho da efetividade:* contribuições para a implementação da LGPD. São Paulo: Thomsom Reuters, 2020. p. 118.

[350] MATTIUZO, M. Discriminação algorítmica: reflexões no contexto da Lei Geral de Proteção de Dados Pessoais. *In*: DONEDA, D.; MENDES, L. S.; CUEVA, R. V. B. (Coords.). *Lei Geral de Proteção de Dados (Lei nº 13.709/2018) — A caminho da efetividade:* contribuições para a implementação da LGPD. São Paulo: Thomsom Reuters, 2020.

de generalizações, conforme explicita: (i) generalizações coerentes (estatisticamente sólidas) e (ii) generalizações problemáticas (estatisticamente falhas). Generalizações problemáticas, afirma, constituem, naturalmente, um problema a ser solucionado. Quanto às generalizações coerentes, por sua vez, cabe analisar se e quando o seu uso deverá ser aceito. Para desenvolver esse raciocínio, a autora faz referência à obra de Frederick Schauer ("Profiles, Probabilities, and Stereotypes"), para quem há três subcategorias de generalizações sólidas: (i) as universais, sendo aquelas que descrevem uma característica comum a todos os integrantes de um determinado grupo; (ii) aquelas que descrevem a maioria de um grupo por meio de uma característica compartilhada por parte significativa de seus integrantes; e (iii) as generalizações comparativas, ou seja, generalizações que levam em com conta uma característica que indivíduos de determinado grupo possuem em maior proporção que indivíduos de outros grupos. Para a autora, as duas últimas subcategorias de generalizações sólidas apresentam maior potencial de questionamento, e trabalhar com essas subcategorias e entender suas diferenças é relevante para a discussão sobre discriminação algorítmica.[351]

Em seguida, destaca que o regime jurídico em torno das discriminações algorítmicas na LGPD parte justamente do art. 6º, IX (princípio da não discriminação), que assume a posição de disposição central na LGPD sobre a matéria. A esse dispositivo somam-se os arts. 20, 11, §5º, e 21, dentre outros.[352]

Ao analisar a redação do princípio da não discriminação, destaca que são vedadas apenas as práticas discriminatórias *abusivas* ou *ilícitas*, concluindo que demais práticas discriminatórias são válidas, porquanto constituem meras generalizações, entendimento coerente com a categorização anteriormente apresentada em seu trabalho.[353] Neste sentido, também destaca ser necessário distinguir as noções de abusividade e ilicitude trazidas pela lei. Conforme esclarece, a *ilicitude* está ligada a uma ideia de proibição legal, como a proibição de coleta de "informações excessivas" para formação de histórico de crédito disposta no art. 3º da Lei do Cadastro Positivo. Um outro tipo de discriminação ilícita de acordo com a LGPD seria aquela baseada em informações estatisticamente incorretas, decorrente de problemas na estruturação do sistema ou de vieses na base de dados. Nestes casos,

[351] Ibidem, p. 119-120.
[352] Ibidem, p. 120.
[353] Ibidem, p. 121.

a verificação do problema pode ser complexa, em razão da própria complexidade e opacidade dos sistemas, pelo que se justifica a necessidade de garantir sua explicabilidade.[354] As discriminações *abusivas*, por sua vez, estão no campo das discriminações estatisticamente corretas, mas que, ainda assim, podem vir a trazer implicações do ponto de vista jurídico, afetando a esfera de direitos de um determinado indivíduo; esclarece que, no caso concreto, o tipo de problema enfrentado — se de ilicitude ou de abusividade (viés dos dados ou disfuncionalidade do algoritmo) — pode-se apresentar de forma híbrida, mas trabalhar com essas categorias é útil para melhor identificar os problemas e endereçá-los.[355]

A autora conclui o artigo afirmando que decisões algorítmicas são essencialmente processos de generalização e discriminação, mas disso não se deve concluir que toda generalização e discriminação seja inaceitável, sendo esse entendimento suportado pelo ordenamento brasileiro. Destaca, ainda, que a principal tarefa interpretativa na implementação da lei é estabelecer distinções entre as noções de ilicitude e abusividade presentes no art. 6º, IX, da LGPD.[356]

Esses aportes teóricos, bem como a distinção e clareza do que viriam ser situações de ilicitude e abusividade, são de extrema importância para o entendimento do escopo do direito à explicação na LGPD, sobretudo para se determinar quando ele seria cabível nas hipóteses em que o emprego de processos discriminatórios possa afetar materialmente a esfera de interesses de um indivíduo, conforme dispõe o art. 20, *caput*, da LGPD.

No contexto de decisões automatizadas, assumem centralidade ainda os princípios da prevenção (art. 6º, VIII) e responsabilização e prestação de contas (art. 6º, X), que, no contexto de crescente reconhecimento da insuficiência de mecanismos de tutela meramente individuais num cenário onde o risco é predominante, surgem como importantes salvaguardas e como mecanismos de tutela estrutural[357], coexistindo com e reforçando os instrumentos de proteção individual já existentes

---

[354] MATTIUZO, M. Discriminação algorítmica: reflexões no contexto da Lei Geral de Proteção de Dados Pessoais. *In*: DONEDA, D.; MENDES, L. S.; CUEVA, R. V. B. (Coords.). *Lei Geral de Proteção de Dados (Lei nº 13.709/2018) — A caminho da efetividade:* contribuições para a implementação da LGPD. São Paulo: Thomsom Reuters, 2020. p. 123.

[355] Ibidem, p. 124.

[356] Ibidem, p. 125.

[357] ALIMONTI, Veridiana. Autodeterminação informacional na LGPD: antecedentes, influências e desafios. *In*: DONEDA, D.; MENDES, L. S.; CUEVA, R. V. B. (Coords.). *Lei Geral de Proteção de Dados (Lei nº 13.709/2018) — A caminho da efetividade:* contribuições para a implementação da LGPD. São Paulo: Thomsom Reuters, 2020. p. 183.

(como os direitos *ad hoc* de oposição e de solicitar revisão de determinada decisão, por exemplo).

Nesse sentido, Alimonti argumenta que há na LGPD um regime de tutela dos dados pessoais que combina ferramentas de controle individuais com elementos de proteção e balanceamento mais estruturais.[358] Como princípios da lei reveladores dessa faceta mais estrutural, a autora aponta os princípios da prevenção (art. 6º, VIII) e da responsabilização e prestação de contas (art. 6º, X). Em razão desses princípios,

> Os controladores devem tanto adotar medidas para prevenir danos em razão do tratamento de dados pessoais quanto para serem capazes de comprovar a observância e o cumprimento das normas de proteção de dados pessoais. Tais medidas se relacionam, ao menos em parte, a ferramentas organizacionais e tecnológicas previstas pela lei em paralelo à "caixa de ferramentas" mais diretamente associada aos poderes individuais de controle ou participação no tratamento de dados pessoais.[359]

Dentre essas ferramentas organizacionais e tecnológicas, estariam o relatório de impacto à proteção de dados pessoais (art. 5º, XVII), a obrigação de garantir que sistemas de tratamento de dados atendam a padrões de segurança, boas práticas de governança e princípios gerais de proteção de dados pessoais (art. 49) e a própria criação da ANPD e do CNPD[360]. A autora conclui seu trabalho afirmando que há muito a visão de controle do fluxo informacional estritamente focada no indivíduo vem sendo insuficiente, o que levou à emergência de uma arquitetura mais complexa de controle, complementar aos direitos e ferramentas individuais, e que a LGPD constitui um exemplo de que tal compreensão mais abrangente e complementar do controle é não apenas possível como também necessária.[361]

Conforme visto no tópico 3.1, o que também abordaremos no capítulo 4 a seguir, também no cenário europeu há o entendimento de que mecanismos meramente individuais são insuficientes para a garantia de um direito pleno à explicação, seja em razão de obstáculo legais, como os direitos de propriedade intelectual, seja em razão de questões de acesso à justiça e assimetria relacional, ou mesmo em razão das limitações de compreensão dos próprios titulares, sendo cada

---

[358] Ibidem, p. 184-188.
[359] Ibidem, p. 189.
[360] Sobre a dimensão institucional da proteção de dados pessoais e dos desafios em relação ao *enforcement* do direito à explicação, cf. Capítulo 4.
[361] ALIMONTI, op. cit., p. 190.

vez mais reconhecido o papel de instrumentos estruturais de proteção, sobretudo mecanismos de regulação *ex ante*, como os relatórios de impacto e as auditorias, mecanismos que se encontram refletidos na LGPD tanto em seus princípios quanto no conjunto de direitos positivados no contexto de decisões automatizados, configurando-se como elementos adicionais aptos a fundamentar a existência de um direito à explicação na LGPD.

Para além do amplo conjunto de princípios, o direito à explicação na LGPD pode ainda ser derivado de alguns direitos a ele relacionados, sobretudo o direito de acesso (arts. 9º, art. 18, I e II, e art. 20, § 1º). Neste sentido, assinalam Souza, Perrone e Magrani:

> A LGPD não seguiu o GDPR no sentido de explicitar no elenco do direito de acesso à informação quais informações o controlador deve prestar e em quais momentos. A lei brasileira não tem artigos similares aos arts. 13(2)(f), 14(2)(g) e 15(1)(h) do GDPR. No entanto, os arts. 9º, I e II, 18, I e II, e 20, § 1º, da LGPD criam uma teia de direitos e obrigações que devem facilitar o acesso à informação.[362]

Para os autores, a LGPD confere aos titulares um direito à informação bastante amplo, com base nos dispositivos acima mencionados. Conforme apontam, as informações devem ser claras e adequadas, sendo possível interpretar que as informações devem servir a um propósito, qual seja, o de possibilitar ao titular o exercício dos seus direitos:

> Quanto à natureza da informação que deve ser apresentada ao titular, a LGPD novamente é similar ao GDPR. Enquanto no GDPR, como vimos, há menção à "informação útil" sobre a "lógica subjacente" e a "importância e as consequências previstas"; na LGPD, há referência a "informações claras e adequadas" e "critérios e procedimentos utilizados". É possível dizer que existem paralelos entre ambas as normas. *"Informações úteis" parece implicar um possível uso para atingir um determinado resultado, que, no caso, seria permitir o exercício dos outros direitos, como de apresentar os seus pontos de vista ou contestar. Similarmente, a expressão "informação adequada" também passa a impressão de que a informação deve ser adequada para atingir um fim, que, no caso, é permitir ao titular exercer seus outros direitos, mormente solicitar a revisão da decisão*.[363] (grifo

[362] SOUZA, C. A.; PERRONE, C.; MAGRANI, E. O direito à explicação entre a experiência europeia e a sua positivação na LGPD. *In*: MENDES, L. S.; DONEDA, D.; SARLET, I. W.; RODRIGUES JÚNIOR, O. L.; BIONI, B. (Orgs.). *Tratado de Proteção de Dados Pessoais*. Rio de Janeiro: Forense, 2021. p. 243-270. p. 263.

[363] SOUZA, C. A.; PERRONE, C.; MAGRANI, E. O direito à explicação entre a experiência europeia e a sua positivação na LGPD. *In*: MENDES, L. S.; DONEDA, D.; SARLET, I. W.; RODRIGUES JÚNIOR,

nosso).

Ainda no que se refere ao direito de acesso, os autores pontuam que a LGPD garante um direito de acesso à informação a qualquer tempo, não reproduzindo a controvérsia sobre o *timing* presente no debate europeu. Além do direito de acesso, estariam ainda relacionados ao direito à explicação o direito de revisão (art. 20, *caput*) e a competência da ANPD para realizar auditorias (art. 20, § 2º), instrumento que serve ao objetivo de proteger os direitos dos titulares, sendo "uma oportunidade de revisão sistêmica", tendo "o potencial de proteger de maneira ampla a população."[364]

Além da possibilidade de reconhecimento do direito à explicação a partir de uma interpretação sistemática e integrativa dos direitos e princípios positivados na LGPD, há também uma visão compartilhada por alguns dos principais autores no debate brasileiro de que o direito à explicação pode ser também entendido como pressuposto para exercício de outros direitos, não apenas aqueles previstos na LGPD, mas também os direitos fundamentais positivados no texto constitucional.

O direito à explicação pode ser entendido, primeiramente, como pressuposto ou condição necessária para o exercício dos direitos de oposição (art. 18, § 2º) e de revisão (art. 20, *caput*), por exemplo. Quanto a isso, assinalam Souza, Perrone e Magrani:

> É justamente porque a lei estabelece certos direitos, como o do pedido de revisão (art. 20, caput), que se torna necessário que em alguma medida os indivíduos sejam esclarecidos sobre os fatores relevantes para a tomada da decisão automatizada. Não diferente do sistema europeu, *não pode existir um exercício efetivo do direito de revisão sem que o indivíduo possa apresentar a sua percepção de como os dados devem ser analisados e de onde podem existir erros, discrepâncias ou mesmo de por que determinado fator não se aplica diretamente a ele ou ela. O direito à explicação é, portanto, no mínimo, um pressuposto para o exercício dos outros direitos, particularmente o de requerer uma revisão*.[365] (grifo nosso).

O. L.; BIONI, B. (Orgs.). *Tratado de Proteção de Dados Pessoais*. Rio de Janeiro: Forense, 2021. p. 243-270. p. 264.

[364] Ibidem, p. 268.

[365] SOUZA, C. A.; PERRONE, C.; MAGRANI, E. O direito à explicação entre a experiência europeia e a sua positivação na LGPD. *In*: MENDES, L. S.; DONEDA, D.; SARLET, I. W.; RODRIGUES JÚNIOR, O. L.; BIONI, B. (Orgs.). *Tratado de Proteção de Dados Pessoais*. Rio de Janeiro: Forense, 2021. p. 243-270. p. 263.

Conforme apontam Silva[366] e Monteiro[367], é possível que decisões automatizadas possam afetar ainda outros direitos, como os direitos fundamentais de acesso à saúde, educação, emprego, o direito à liberdade, à cidadania, dentre outros. Essa visão encontra suporte ainda em Mulholland e Frahjof. Conforme argumentam as autoras, em determinadas circunstâncias, o direito à explicação pode ser entendido como condição para a garantia de outros direitos fundamentais, sobretudo quando determinada decisão automatizada estiver mediando o acesso a um bem jurídico tutelado constitucionalmente, como é o caso do acesso à educação via financiamento estudantil, do acesso ao crédito para fins de aquisição de moradia, etc. Afirmam o seguinte:

> Considerando que cada vez mais estas aplicações têm um forte impacto em áreas sensíveis da sociedade, tais como o uso de dados para o desenvolvimento de ajuda humanitária, auxílio no diagnóstico médico correto ou proporcionar racionalidade a decisões judiciais (ALMEIDA; DONEDA, 2019, p. 143), é certo o potencial impacto que estas decisões automatizadas poderão causar aos direitos individuais e coletivos (artigo 5º da CFRB) dos titulares da dados, mas também aos seus direitos sociais (artigo 6º da CRFB). Em casos envolvendo decisões automatizadas que irão conceder ou negar determinado bem jurídico, é essencial avaliar a natureza de tal bem. *Caso este bem vá realizar funções sociais constitucionalmente asseguradas (tais como a concessão de financiamento de crédito estudantil ou a obtenção de crédito para aquisição de moradia), concretizando direitos essenciais, a compreensão e o exercício de um direito à explicação serão primordiais. Por isso, a definição do que consiste ser o dever do controlador em atribuir uma explicação ao algoritmo de tomada de decisão vai influenciar não apenas a capacidade de compreensão do titular de reconhecer se houve ou não uma discriminação, mas vai viabilizar o exercício de outros direitos fundamentais*, tendo como objetivo construir uma sociedade livre, justa e solidária (artigo 3º, I, da CRFB).[368] (grifo nosso).

Como se pode observar, há não apenas uma significativa convergência acerca da existência de elementos aptos a fundamentar a existência de um direito à explicação na LGPD, como também uma diversidade de formas de fundamentá-lo,

---

366 SILVA, P. R. Os direitos dos titulares de dados. *In*: MULHOLLAND, Caitlin (Org.). *A LGPD e o novo marco normativo no Brasil*. Porto Alegre: Arquipélago, 2020. p. 210.

367 MONTEIRO, R. L. *Existe um direito à explicação na Lei Geral de Proteção de Dados do Brasil?* Rio de Janeiro: Instituto Igarapé, dez. 2018. Artigo Estratégico nº 39. Disponível em: https://igarape.org.br/wp-content/uploads/2018/12/Existe-um-direito-a-explicacao-na-Lei-Geral-de-Protecao-de-Dados-no-Brasil.pdf. Acesso em: 21 jun. 2021. p. 2-4.

368 MULHOLLAND, C.; FRAJHOF, I. Z. Inteligência Artificial e a Lei Geral 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. (Coords.). *Inteligência Artificial e Direito:* Ética, Regulação e Responsabilidade. 2. ed. São Paulo: Revista dos Tribunais, 2020.

seja a partir da derivação de disposições principiológicas da lei, seja a partir de uma análise integrativa dos direitos nela previstos, seja a partir do entendimento do direito à explicação como pressuposto para o exercício de outros direitos positivados na LGPD e na ordem constitucional.

Para além desse esforço de fundamentação, parte do debate brasileiro tem se concentrado em analisar criticamente algumas das disposições da LGPD relacionadas ao direito à explicação, explicitando suas especificidades em relação ao regulamento europeu. O tópico a seguir buscará sistematizar essas considerações, tendo como objetivo analisar em que medida a LGPD dá abertura ao reconhecimento de um direito à explicação mais restrito ou mais amplo em relação àquele que pode ser extraído do regulamento europeu.

3.2.2.3 O regime jurídico aplicável às decisões automatizadas: colocando em perspectiva LGPD e GDPR

Apesar da existência de inúmeros pontos de intersecção com o regulamento europeu, há diversas peculiaridades no regime jurídico das decisões automatizadas na LGPD que merecem destaque. A depender do dispositivo e das características do regime consideradas, bem como das interpretações atribuídas a cada dispositivo, é possível reconhecer ora um regime de escopo mais amplo, ora um regime de escopo mais restrito em relação àquele desenhado na GDPR.

Analisaremos essas convergências e divergências em dois eixos: primeiramente, analisaremos as especificidades dos direitos existentes no contexto de decisões automatizadas, apontando em quais pontos os direitos consagrados na LGPD seriam mais amplos ou mais restritos em relação aos seus correspondentes na GDPR; em seguida, retomaremos a discussão acerca do alcance do conceito de dado pessoal na LGPD, apresentada no Capítulo 1, que, como vimos, refletiu em seu texto a teoria consequencialista, o que confere ao conceito de dado pessoal na LGPD um possível escopo maior em relação àquele adotado na GDPR, o que traz consequências para o alcance de um direito à explicação no cenário brasileiro.

Um primeiro paralelo possível de ser traçado diz respeito à natureza jurídica dos arts. 22(1) da GDPR e art. 20, *caput*, da LGPD. Como abordado no item 3.1, o art. 22(1) da GDPR estabelece como regra geral a proibição de decisões baseadas

unicamente no processamento automatizado de dados[369]. Como visto, essa regra geral comporta exceções, quais sejam: (i) quando a decisão automatizada for necessária para a performance de um contrato; (ii) quando autorizada pelo direito interno de um Estado-Membro, que deverá dispor ainda sobre as salvaguardas adequadas para proteger os direitos, liberdades e legítimos interesses dos titulares, e (iii) quando autorizada mediante consentimento explícito do titular. Na primeira e na terceira hipóteses, o controlador deverá adotar medidas adequadas para salvaguardar os direitos, liberdades e legítimos interesses do titular, devendo garantir, ao menos, a revisão por pessoa natural, bem como garantir ao titular o direito de expressar seu ponto de vista em relação à decisão e de contestá-la. A regra do art. 20 da LGPD, por sua vez, possui natureza jurídica distinta, caracterizando-se como um direito de oposição. Em outras palavras, isso implica dizer que, na LGPD, as decisões automatizadas não estão, via de regra, proibidas, devendo, todavia, observarem determinadas salvaguardas legais, que incluem, mas não se limitam, ao direito de revisão.[370] Nas palavras de Mulholland e Frahjof: "Isto é, no ordenamento europeu, a norma tem uma *natureza proibitiva*, vedando a tomada de decisões totalmente automatizada, enquanto no ordenamento brasileiro, a norma tem *natureza atributiva de direitos*."[371] (grifo nosso).

Em segundo lugar, cabe observar que a GDPR previu bases legais específicas para o processamento automatizado de dados pessoais (execução do contrato, quando autorizado mediante o direito interno de um Estado-Membro e com base no consentimento do titular). De igual modo, é possível afirmar que os direitos relativos a decisões automatizadas no contexto do regulamento são também limitados em relação a essas bases legais. A LGPD, todavia, não estabeleceu qualquer limitação em relação à base legal a ser empregada para legitimar o processamento automatizado de dados pessoais, estando, a princípio, autorizada a atribuição de quaisquer das hipóteses autorizadoras previstas nos arts. 7º e 11.

Um terceiro ponto diz respeito ao direito de revisão de decisões automatizadas

---

[369] EUROPEAN DATA PROTECTION BOARD. *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 679/2016*. Bruxelas: European Commission, 2016. Disponível em: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053. Acesso em: 21 jun. 2021.

[370] SOUZA, C. A.; PERRONE, C.; MAGRANI, E. O direito à explicação entre a experiência europeia e a sua positivação na LGPD. *In*: MENDES, L. S.; DONEDA, D.; SARLET, I. W.; RODRIGUES JÚNIOR, O. L.; BIONI, B. (Orgs.). *Tratado de Proteção de Dados Pessoais*. Rio de Janeiro: Forense, 2021. p. 243-270. p. 261.

[371] MULHOLLAND; FRAJHOF, op. cit.

na GDPR e na LGPD. O Regulamento europeu, a partir de uma leitura do Considerando 71[372] e das *Guidelines on Automated-individual decision-making and Profiling for the purposes of Regulation 2016/679*, contempla um direito à revisão por pessoa natural, cuja intervenção deverá ser significativa, e não meramente *pro forma*[373], conforme discutido no item 3.1. Na lei brasileira, todavia, não há a previsão expressa de que a revisão de decisões automatizadas se dê mediante intervenção de uma pessoa natural. Essa previsão havia sido incluída na forma de um § 3º ao art. 20, adicionado pelo Congresso Nacional quando da conversão da Medida Provisória nº 869, de 2018, tendo sido vetada pelo então presidente Michel Temer e posteriormente por Jair Bolsonaro.[374] O texto da LGPD, portanto, da forma como está hoje posto, não prevê expressamente a garantia de revisão por pessoa natural, mas não a veda. Contudo, a interpretação que tem sido ventilada pela maior parte dos(das) autores(as) que se debruçaram sobre o assunto até o momento, tem sido no sentido de interpretar o direito à revisão de decisões automatizadas de forma ampliativa, sob pena de esvaziamento da sua tutela, em consonância com os princípios e com a própria racionalidade adotada pela LGPD.

Neste sentido, conforme assinalam Bioni e Mendes, ainda que o termo "humana" tenha sido suprimido da redação do caput do art. 20, por uma interpretação sistemática da LGPD, a intervenção humana continua se fazendo necessária em alguma fase do processo de revisão. Caso essa interpretação prevaleça no direito brasileiro, haveria então uma aproximação dos dois regimes quanto a esse aspecto[375],

---

[372] Considerando 71: ""which should include specific information to the data subject and the right to obtain human intervention, to express his or her point of view, to obtain an explanation of the decision reached after such assessment and to challenge the decision. Such measure should not concern a child." (UNIÃO EUROPEIA. *Regulamento Geral sobre a Proteção de Dados — Considerando 71*. Disponível em: https://gdpr-text.com/pt/read/recital-71/. Acesso em: 21 jun. 2021). (grifo nosso).

[373] "Human intervention is a key element. Any review must be carried out by someone who has the appropriate authority and capability to change the decision. The reviewer should undertake a thorough assessment of all the relevant data, including any additional information provided by the data subject." Cf. EUROPEAN DATA PROTECTION BOARD. *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 679/2016*. Bruxelas: European Commission, 2016. Disponível em: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053. Acesso em: 21 jun. 2021.

[374] Para uma 'genealogia' do direito à revisão de decisões automatizadas na LGPD, cf. DATA PRIVACY BRASIL. Memória da LGPD: como a lei mudou desde 2010? *Observatório da Privacidade*, 2019. Disponível em: https://www.observatorioprivacidade.com.br/memoria/como-a-lei-mudou-desde-2010/. Acesso em: 11 abr. 2021; e SOUZA, C. A.; PERRONE, C.; MAGRANI, E. O direito à explicação entre a experiência europeia e a sua positivação na LGPD. *In*: MENDES, L. S.; DONEDA, D.; SARLET, I. W.; RODRIGUES JÚNIOR, O. L.; BIONI, B. (Orgs.). *Tratado de Proteção de Dados Pessoais*. Rio de Janeiro: Forense, 2021. p. 243-270. p. 260-261; 266.

[375] BIONI, B; MENDES, L. S. Regulamento Europeu de Proteção de Dados Pessoais e a Lei Geral

o que é altamente desejável no contexto de uma eventual decisão de adequação. Alimonti, por sua vez, ao comentar a retirada da previsão expressa de revisão por pessoa natural, afirma que "[...] é de se discutir a efetividade dessa garantia caso as revisões sejam, como regra, realizadas pelos mesmos sistemas que causaram o questionamento, com seus possíveis vieses ou falhas."[376]

Sobre esse aspecto, Silva defende a posição de que, apesar da supressão do termo "humana", a possibilidade de revisão humana de decisões automatizadas continua a existir, estando ausentes apenas as condições detalhadas da revisão prevista no art. 20.[377] A autora argumenta que, ao se retirar a expressão "humana", abre-se a possibilidade de que uma revisão de decisão automatizada seja realizada por um sistema também automatizado, e não por um humano, "[...] prejudicando a transparência e a concretização de um direito à explicação consistente."[378]

Souza, Perrone e Magrani defendem que, de modo a garantir um efetivo direito de revisão no contexto de decisões automatizadas, deve ser considerada como boa prática a revisão por uma pessoa natural:

> A redação atual da lei não demanda a revisão por pessoa natural. No entanto, deve-se entender que, para garantir o pleno exercício do direito de revisão, este deve ser efetivo e permitir que se possa chegar a conclusões diferentes das apresentadas pela decisão automatizada original. Desse modo, deve-se considerar que a revisão por uma pessoa natural é uma prática recomendável, sempre e quando seja possível e pertinente para os fins aqui debatidos.[379]

Ainda, quanto ao direito de revisão, Mulholland e Frajhof chamam a atenção para o fato de que "[...] a lei autoriza o pedido de revisão, mas isto não significa que, após a análise pelo controlador, o resultado final necessariamente será alterado."[380]

---

brasileira de Proteção de Dados: mapeando convergências na direção de um nível de equivalência. *In*: TEPEDINO, G.; FRAZÃO, A.; OLIVA, M. D. (Coords.). *Lei Geral de Proteção de Dados Pessoais e suas repercussões no Direito Brasileiro*. 2. ed. São Paulo: Thomson Reuters, 2020. p. 803.

[376] ALIMONTI, V. Autodeterminação informacional na LGPD: antecedentes, influências e desafios. *In*: DONEDA, D.; MENDES, L. S.; CUEVA, R. V. B. (Coords.). *Lei Geral de Proteção de Dados (Lei nº 13.709/2018) — A caminho da efetividade:* contribuições para a implementação da LGPD. São Paulo: Thomsom Reuters, 2020. p. 189.

[377] SILVA, P. R. Os direitos dos titulares de dados. *In*: MULHOLLAND, Caitlin (Org.). *A LGPD e o novo marco normativo no Brasil*. Porto Alegre: Arquipélago, 2020. p. 211.

[378] Ibidem, p. 212.

[379] SOUZA, C. A.; PERRONE, C.; MAGRANI, E. O direito à explicação entre a experiência europeia e a sua positivação na LGPD. *In*: MENDES, L. S.; DONEDA, D.; SARLET, I. W.; RODRIGUES JÚNIOR, O. L.; BIONI, B. (Orgs.). *Tratado de Proteção de Dados Pessoais*. Rio de Janeiro: Forense, 2021. p. 243-270. p. 267.

[380] MULHOLLAND, C.; FRAJHOF, I. Z. Inteligência Artificial e a Lei Geral de Proteção de Dados

Uma outra importante discussão no contexto do direito à revisão de decisões automatizadas diz respeito aos efeitos materiais necessários para caracterizar determinada decisão automatizada como juridicamente relevante.

Nos termos do art. 22(1) da GDPR, o titular dos dados possui o direito de não estar sujeito a uma decisão automatizada, incluindo a criação de perfis, que produza efeitos jurídicos a seu respeito ou que o afete de forma semelhante. Como se vê, não é qualquer decisão automatizada que atrai a proteção do artigo, mas apenas aquelas capazes de produzir efeitos na esfera jurídica do titular ou de afetá-lo de forma similar. As *Guidelines on Automated-individual decision-making and Profiling for the purposes of Regulation 2016/679* oferecem importantes diretrizes para a melhor compreensão do que viriam a ser "efeitos jurídicos" ou "efeitos semelhantes". De acordo com o WP29, a GDPR não chega a esclarecer o que seriam esses efeitos, mas a redação adotada parece sugerir que apenas impactos ou efeitos significativos estariam contemplados pelo artigo. De acordo com as diretrizes, um "efeito jurídico" requer que a decisão seja capaz de afetar a esfera de direitos do indivíduo, seu *status* legal ou seus direitos no âmbito de um contrato. Exemplos de efeitos dessa ordem incluem decisões automatizadas que possam resultar, por exemplo, no cancelamento de um contrato, na concessão ou negativa de determinado benefício social ou negativa de cidadania.[381] Ainda que determinado efeito não possa ser caracterizado como um "efeito jurídico", ele ainda poderá ser alcançado pelo escopo do art. 22 se possuir efeitos equivalentes ou for igualmente relevante em seu impacto. Neste sentido, decisões que carreguem o potencial de: (i) afetar significativamente as circunstâncias, comportamento ou escolhas dos indivíduos envolvidos; (ii) ter um impacto prolongado ou permanente na pessoa em causa; ou (iii) em sua forma mais extrema, levar à exclusão ou discriminação de indivíduos, podem ser entendidas como similarmente relevantes. O documento deixa claro que é difícil precisar o que seria considerado suficientemente significativo, mas apresenta alguns exemplos de situações que entrariam nesse escopo: (i) decisões que afetam as circunstâncias financeiras de alguém, como sua elegibilidade ao crédito; (ii) decisões que afetam o acesso de

---

Pessoais: breves anotações sobre o direito à explicação perante a tomada de decisões por meio de machine learning. *In*: FRAZÃO, A.; MULHOLLAND, C. (Coords.). *Inteligência Artificial e Direito:* Ética, Regulação e Responsabilidade. 2. ed. São Paulo: Revista dos Tribunais, 2020.

[381] EUROPEAN DATA PROTECTION BOARD. *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 679/2016*. Bruxelas: European Commission, 2016. Disponível em: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053. Acesso em: 21 jun. 2021. p. 21.

alguém aos serviços de saúde; (iii) decisões que neguem o acesso a uma oportunidade de emprego ou que coloquem o indivíduo em uma séria desvantagem; e (iv) decisões que afetam o acesso de alguém à educação, como o processo de admissão em uma universidade.

Como visto acima, a GDPR estabelece uma forte barreira a ser superada para que se reconheça determinada decisão automatizada como dotada de efeitos jurídicos relevantes ou similarmente significativos, apta, portanto, a atrair a proteção do art. 22. No Brasil, esse limiar é menor: o direito à revisão existe quando o procedimento automatizado for capaz de impactar os interesses do titular. “Interesses”, tal como positivado no *caput* do art. 20 da LGPD, é uma barreira muito mais fácil de ser ultrapassada do que o limite dos interesses jurídicos materiais exigidos pela GDPR. É este o entendimento, por exemplo, de Hosni e Martins, que apontam para a existência de uma noção de *interesse* muito mais ampla na LGPD do que aquela contida na GDPR:

> Destaca-se que o termo “interesse” dá maior abrangência a essa norma, não sendo necessária a verificação de uma violação de um direito específico para que o art. 20 possa ser invocado. O simples fato de uma decisão totalmente automatizada afetar interesses do titular (o que também inclui ameaças a direitos) já é o suficiente para sua aplicação. Portanto, ao se diferenciar da GDPR, que restringe a incidência de seu art. 22 para decisões automatizadas que produzam efeitos legais ou, de maneira similar, significativamente afete o titular, a LGPD torna possível uma tutela preventiva por parte do titular, antes mesmo de se caracterizar um dano efetivo.[382]

Por fim, ainda quanto ao art. 20 da LGDP, Mulholland e Frahjof fazem uma importante ressalva quanto à competência da Autoridade Nacional de Proteção de Dados para realizar auditorias em sistemas algorítmicos na hipótese de o controlador se recusar, alegando segredo comercial ou industrial, a fornecer informações claras e adequadas a respeito dos critérios e procedimentos empregados na decisão automatizada. A redação do art. 20, § 2º, expressamente prevê que “a autoridade nacional *poderá* realizar auditoria”[383] (grifo nosso), o que, para as autoras, denota a

[382] 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*, v. 1, n. 2, dez. 2020, p. 90. Disponível em: https://revista.internetlab.org.br/wp-content/uploads/2020/12/Tomada-de-Decisa%CC%83o-Automatizada.pdf. Acesso em: 11 abr. 2021. p. 90.

[383] BRASIL*. Lei nº 13.709, de 14 de agosto de 2018*. Lei Geral de Proteção de Dados Pessoais (LGPD).

existência de uma margem de discricionariedade da ANPD para exercer ou não sua competência fiscalizatória. Tal como está posta, esta redação pode dar margem a comportamentos indesejados e oportunistas por parte do ente regulado:

> A segunda reconhece, à primeira vista, a discricionariedade da autoridade nacional para realizar a auditoria apenas quando o controlador se negar a fornecer as informações elencadas no parágrafo primeiro. A existência desta condição pode dar margem para que o controlador se negue a exercer a explicação com base em uma simples alegação de que seu código estaria protegido pelo segredo comercial ou industrial, pois sabe que a atuação da autoridade em auditar seu algoritmo será optativa. Por sua vez, caso a explicação concedida ao titular não seja suficientemente clara – impedindo que a pessoa seja capaz de compreender se houve ou não um tratamento discriminatório, ou o motivo pelo qual o algoritmo decidiu de uma maneira e não de outra –, parece que o titular de dados terá menos garantias do que se o controlador de dados tivesse meramente alegado a proteção do sigilo do seu algoritmo.[384]

Como afirmamos no início desta subseção, é ainda possível argumentar que o próprio conceito de dado pessoal adotado pela LGPD é mais abrangente do que aquele previsto na GDPR. Conforme abordamos no tópico 1.1.1 (“O que são ‘dados pessoais’? Uma visão consequencialista e a necessidade de um novo conceito”), a LGPD não apenas se filia à abordagem expansionista do conceito de dado pessoal como vai além, inaugurando a teoria consequencialista, trazendo relevantes consequências práticas ao expandir o conceito de dado pessoal para o uso de todo e qualquer tipo de dado que possa ter impacto sobre o indivíduo ou uma coletividade, ampliando de forma significativa o escopo de aplicação da LGPD. Nesse sentido, ao expandir de forma significativa o conceito de dado pessoal, a LGPD permite que uma gama maior de processos automatizados sejam alcançados pelo seu escopo de aplicação, abrindo margem, ao menos quanto a este aspecto, ao reconhecimento de um direito à explicação de contornos mais alargados do que aquele possível de ser extraído de uma interpretação da GDPR. Em suma, se a teoria consequencialista inclui tipos de dados em que o seu tratamento impacta indivíduos, significa que processos automatizados que utilizam tipos de dados que não se encaixam no

---

Brasília, DF: Presidência da República, 14 ago. 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015- 2018/2018/lei/L13709.htm. Acesso em: 07. nov. 2020.

[384] MULHOLLAND, C.; FRAJHOF, I. Z. Inteligência Artificial e a Lei Geral 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. (Coords.). *Inteligência Artificial e Direito:* Ética, Regulação e Responsabilidade. 2. ed. São Paulo: Revista dos Tribunais, 2020.

conceito reducionista e expansionista podem estar incluídos no conceito consequencialista e, portanto, dentro do escopo da LGPD (ideia de *single out*).

Como visto, apesar de a LGPD oferecer elementos para o reconhecimento de um direito à explicação, ainda há incerteza quanto ao escopo e ao exato desenho deste direito no ordenamento brasileiro. Conforme apontam Mulholland e Frajhof:

> Em relação aos pressupostos para a identificação de um direito à explicação, o GDPR e a LGPD relegam à doutrina o encargo de sua delimitação. Assim, para a compreensão dos limites de um direito à explicação, é necessário precisar "o que é uma decisão totalmente automatizada, que tipos de decisão automatizada afetam a esfera jurídica dos titulares de dados e qual é o grau de transparência e explicação que será exigível em situações assim.[385]

Como apresentado anteriormente, a doutrina tem se debruçado sobre o texto da LGPD na tentativa de interpretá-lo e dele extrair elementos para o reconhecimento de um direito à explicação, mas há ainda uma escassez de regulamentação e acúmulo jurisprudencial que comprometem a exata compreensão de alguns dispositivos.

Nesse aspecto, Mulholland e Frajhof apontam que,

> Considerando a indefinição sobre o assunto, será importante acompanhar o desenvolvimento da interpretação do direito à explicação pelas autoridades de proteção de dados pessoais brasileira e europeia, assim como pelos próprios tribunais e cortes de justiça.[386]

Com relação ao exposto, tem-se reconhecido a importância e urgência da atuação da ANPD, bem como de outros atores, no esclarecimento dessas controvérsias e no enfrentamento das incertezas em torno do direito à explicação na LGPD:

> Os limites e potencialidades desse direito dependerão, então, da atuação de muitos agentes. Em primeiro lugar, controladores de dados poderão desenvolver recomendações sobre quando e como a intervenção humana se torna desejável, até como medida de incremento da confiança por parte do titular de dados. De outro lado, a autoridade de proteção de dados também poderá se pronunciar sobre o tema, traçando diretrizes para sua implementação. Esse quadro não estaria completo sem a menção à atuação de outras entidades, como os Procons, o Ministério Público e o próprio Poder

[385] MULHOLLAND, C.; FRAJHOF, I. Z. Inteligência Artificial e a Lei Geral 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. (Coords.). *Inteligência Artificial e Direito:* Ética, Regulação e Responsabilidade. 2. ed. São Paulo: Revista dos Tribunais, 2020.
[386] Ibidem.

> Judiciário, que nas atividades que lhes são próprias ajudarão a definir os contornos do direito à explicação.[387]

A exata compreensão da extensão de um direito à explicação no ordenamento brasileiro é também dependente do desenvolvimento da discussão no cenário internacional, como mencionamos no início desta seção. De acordo com Mulholland e Frajhof:

> Embora se reconheça a existência de diferenças entre o direito à explicação previsto no GDPR e na LGPD, as lições aprendidas no âmbito da Europa, que possui ampla experiência na regulação das atividades de tratamento de dados pessoais, serão importantes para auxiliar nas definições deste direito no contexto brasileiro, que deve levar em consideração o correto balanceamento entre os interesses dos titulares de dados e o direito à livre iniciativa dos controladores de dados. As dúvidas sobre como implementar a explicação e suas complexidades técnicas, sobre qual conteúdo deverá ser fornecido pelo controlador e em que momento o titular de dados poderá pleitear tal direito ainda precisarão ser amadurecidas e respondidas. O que se pode concluir, até o momento, é que as previsões regulando este contexto nos permite reconhecer a relevância do assunto, e a preocupação dos legisladores, cientes dos desafios que os sistemas de IA aplicados a algoritmos de tomada de decisão colocam, de preverem a possibilidade de que o titular de dados não fique sujeito a decisões totalmente automatizadas, e que tenha o direito de requerer uma explicação sobre esta decisão. [...] Somado a isto, deve-se acompanhar e adicionar à discussão as lições aprendidas, e as que ainda irão surgir, com a controvérsia que se dá no âmbito do GDPR sobre a existência ou não de tal direito.[388] (grifo nosso).

No Capítulo a seguir, no item 4.3, buscaremos discutir, apesar da escassez de precedentes no cenário brasileiro e mesmo de pronunciamentos ou iniciativas da ANPD até o presente momento, como essas questões vêm sendo endereçadas pelas cortes nacionais e os desafios a serem enfrentados pela própria Autoridade Nacional de Proteção de Dados no que diz respeito ao direito à explicação.

---

[387] SOUZA, C. A.; PERRONE, C.; MAGRANI, E. O direito à explicação entre a experiência europeia e a sua positivação na LGPD. *In*: MENDES, L. S.; DONEDA, D.; SARLET, I. W.; RODRIGUES JÚNIOR, O. L.; BIONI, B. (Orgs.). *Tratado de Proteção de Dados Pessoais*. Rio de Janeiro: Forense, 2021. p. 243-270. p. 268.

[388] MULHOLLAND, C.; FRAJHOF, I. Z. Inteligência Artificial e a Lei Geral 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. (Coords.). *Inteligência Artificial e Direito:* Ética, Regulação e Responsabilidade. 2. ed. São Paulo: Revista dos Tribunais, 2020.

3.3 SÍNTESE DO CAPÍTULO

Neste capítulo apresentamos os aspectos regulatórios do direito à explicação. A partir da discussão das possibilidades de afirmação da existência de um direito apresentada no capítulo anterior, desenvolveu-se de forma mais detalhada os sistemas jurídicos de proteção de dados e as possibilidades de interpretação da transparência e da governança algorítmica. Esse aspecto, como demonstramos, pode ser observado por pelo menos duas óticas distintas, quais sejam, a ótica do devido processo informacional presente no debate estadunidense e sob a ótica do direito à explicação como derivado das obrigações de transparência da proteção de dados, próprios dos modelos Europeu e Brasileiro.

A experiência norte-americana permite compreender como a governança algorítmica pode ocorrer mesmo na ausência de uma autoridade nacional e de como os processos tecnológicos podem ser compreendidos em termos de justiça procedimental, levando em consideração os riscos emergentes das tecnologias de *big data* a partir da noção de devido processo informacional.

O debate sobre a regulação e a jurisprudência europeia permitiu compreender o papel do direito à explicação no regime de *accountability* previsto naquele modelo compreensivo de regulação de dados. A partir da apresentação dos precedentes legislativos e da interpretação dos tribunais brasileiros sobre os *scores* de crédito, foi possível definir como o direito à explicação pode ser compreendido no ordenamento brasileiro. Em comparação com a realidade europeia, como apresentou-se a partir de uma leitura de diversos autores relevantes no debate nacional, é possível afirmar que a LGPD fornece um corpo regulatório que garante um direito à explicação de forma mais consistente do que o GDPR.

A partir dos apontamentos dos autores estrangeiros e brasileiros sobre o direito à explicação no regramento Europeu, bem como da LGPD, foi possível apresentar algumas limitações ao direito à explicação, como a dificuldade de interpretação de alguns conceitos previsto no regulamento, que carecem de uma elaboração dos legisladores e das autoridades nacionais, bem como a dificuldade de sua aplicação prática em relação aos “interesses de terceiros”, entendidos como aspectos de propriedade industrial, intelectual e os segredos de negócio. Outra limitação apontada diz respeito às discussões sobre as tecnologias de *machine learning* e suas implicações para o debate da transparência algorítmica, visto que algumas dessas

tecnologias representam desafios cognitivos relevantes. Essas limitações serão exploradas no próximo capítulo, de forma mais detida, enquanto limitações legislativas, cognitivas e institucionais, como forma de definir contornos mais precisos do alcance do direito à explicação.

4 LIMITES TEÓRICOS E PRÁTICOS PARA A EFETIVAÇÃO DO DIREITO À EXPLICAÇÃO

O direito à explicação pode ser visto como um corolário do princípio da transparência, com vistas ao exercício da autodeterminação informacional. Como discutido no Capítulo 2, diferente da mera obrigação de transparência, a explicação envolve uma ação de interação, que visa a um objetivo específico, qual seja, garantir a efetiva compreensão, por parte do titular de dados, dos elementos lógicos, informacionais, operacionais e dos fundamentos constitutivos do processo de decisão automatizada. Esse direito permite que o titular de dados seja capaz de tomar medidas para verificar a correção dos dados ou questionar os pressupostos da decisão.

A necessidade dessa explicação se justifica, do ponto de vista prático, por uma característica própria de alguns sistemas computacionais. As aplicações são operadas por códigos em linguagem computacional em uma dimensão que não está, na maioria das vezes, acessível ao conhecimento imediato do sujeito. A bibliografia adotou o termo opacidade, isto é, a qualidade oposta da transparência, para definir essa característica dos algoritmos e sistemas de decisões automatizadas em geral.

O direito à explicação, como vimos, visa superar essa opacidade para garantir o exercício da autodeterminação informacional. Portanto, é mister compreender as causas dessa opacidade, uma vez que ela impõe diferentes desafios teóricos e jurídicos para a garantia do direito à explicação. Em um trabalho crucial sobre o problema, Jenna Burrel[389] apresenta três modalidades distintas de opacidade. A primeira delas diz respeito à opacidade como uma estratégia comercial, que envolve problemas relacionados à função econômica dos algoritmos. Esse é o problema que será discutido no item 4.1 deste capítulo. Para a discussão do direito à explicação, será importante ainda apontar alguns desafios relacionados ao conceito de dados pessoais nas legislações de privacidade e proteção de dados.

Outra causa da opacidade apontada pela autora diz respeito ao nível de conhecimento técnico dos sujeitos sobre sistemas automatizados. Pessoas comuns não são letradas e treinadas na linguagem computacional, de forma que o funcionamento dos *softwares* é algo de difícil compreensão. E, por último, segundo

[389] BURRELL, J. *How the Machine "Thinks:" Understanding Opacity in Machine Learning Algorithms*. Rochester, NY: Social Science Research Network, 2015. Disponível em: https://papers.ssrn.com/abstract=2660674. Acesso em: 27 maio. 2020.

ela, há uma limitação própria da natureza de alguns sistemas computacionais, principalmente no campo do aprendizado de máquina (*machine learning*). Certas aplicações realizam operações matemáticas e procedimentos tão complexos que sua compreensão se torna um desafio até mesmo para os técnicos que as desenvolveram. São estes os desafios cognitivos, que exploraremos no item 4.2.

Além dessas questões teóricas, o direito à explicação depende da capacidade de *enforcement* das autoridades (agências nacionais, poder judiciário etc.) para fazer valer as obrigações de transparência e os demais direitos dos titulares. Contudo, as entidades regulatórias encontram diversas dificuldades, técnicas ou administrativas, para exercer essa função. Dessa forma, discutiremos no item 4.3 os desafios institucionais para o exercício do direito à explicação.

4.1 DESAFIOS LEGISLATIVOS: O MANTO DO SEGREDO DE NEGÓCIO E A LIMITAÇÃO DO CONCEITO DE DADOS PESSOAIS

A primeira das limitações ao direito à explicação advém dos limites impostos pela própria legislação. Como apresentado anteriormente, o direito à explicação é resultado de direitos da personalidade, sobretudo o direito à autodeterminação informativa. Como se pode imaginar, ele não tem a pretensão de se sobrepor completamente às demais proteções estabelecidas pelo ordenamento jurídico.

Para compreender essas limitações, é importante mapear as outras esferas jurídicas que se impõem para tutelar direitos concorrentes. Para tanto, valemo-nos das definições e do enquadramento conceitual dado no item 1.1.1 deste trabalho, que explica com maior detalhe como se opera um algoritmo.

Um algoritmo é uma série de procedimentos que visa transformar uma entrada de dados (*input*) em um resultado (*output*), a partir de uma série de operações determinadas ou determináveis. Explicamos que há nesse processo três elementos principais, intensamente relacionados entre si, mas ainda assim delimitados: no nível mais elementar, temos os (i) dados que são submetidos a um (ii) algoritmo, cuja operação servirá a uma (iii) implementação na sociedade[390].

Trabalhando uma correspondência com esses elementos de análise, entendemos que a explicação é um direito que surge no primeiro elemento dessa

[390] MITTELSTADT, B. D. *et al.* The ethics of algorithms: Mapping the debate. *Big Data & Society*, [*S. l.*], v. 3, n. 2, 2016. Disponível em: https://doi.org/10.1177/2053951716679679. Acesso em: 20 maio 2020.

cadeia, apoiada na própria existência do dado pessoal. Como a norma incide sobre o substrato sobre o qual opera toda essa cadeia, e se valendo de proteções jurídicas fortes ligadas ao direito da personalidade, as garantias que ela carrega acabam por refletir também nos processos e implementações. Isso é coerente com a proteção ampla da personalidade e no regime de controle de fluxos de dados, sobretudo no que tange à autodeterminação, e não apenas uma questão de confidencialidade de informações[391].

Contudo, ao percorrer essa cadeia, ela encontra outros direitos que incidem sobre os sistemas informacionais dentro dos quais os dados são tratados. Nesse caso, a principal barreira vem da legislação de Propriedade Intelectual. O regime de proteção de algoritmos, no Brasil, exige confidencialidade e gera uma força antagônica à explicação, que é um mecanismo de transparência.

4.1.1 O algoritmo como um segredo e uma estratégia comercial

O algoritmo, puro e simples, é um modelo matemático que não goza de uma proteção de exclusividade direta pela legislação brasileira. Por isso, caso ele seja tornado público, nada impediria que a empresa que desenvolveu o algoritmo específico visse seus concorrentes usando a mesma receita. Além do mais, os procedimentos de um algoritmo não são diretamente ligados ao procedimento lógico utilizado, que nós humanos representamos com o código, mas que em essência são complexas cadeias de operações eletrônicas. Por isso, os elementos conceituais dos algoritmos podem ser replicados sem necessariamente se copiar literalmente o mesmo texto (código), tornando uma tutela eficaz ainda mais difícil.

Portanto, o regime de segredo industrial se tornou por excelência o meio das organizações conservarem a sua competitividade e capitalizarem em cima dos investimentos que fizeram desenvolvendo novos serviços. O algoritmo traduz-se em mais do que um “tempero secreto” de uma receita, mas também como uma estratégia comercial da empresa.

Por exemplo, no famoso caso da varejista americana Target, os analistas de dados da empresa desenvolveram métodos para identificar clientes grávidas, para

[391] BIONI, B. R. *Proteção de dados pessoais:* a função e os limites do consentimento. 1. ed. Rio de Janeiro: Forense, 2019.

oferecer produtos mais ajustados para esse segmento[392]. Ter um algoritmo que faz essa avaliação não é meramente uma questão de técnica, mas também uma questão de estratégia comercial. A opção por atrair mulheres grávidas a fazerem compras em suas lojas, por saberem que elas acabam atraindo toda uma família de consumidores, foi uma decisão de mercado que deu à Target uma vantagem competitiva. Essa decisão veio de uma visão de negócio de quem trabalhou a estratégia comercial da empresa, e os algoritmos e a coleta de dados foram ferramentas eficazes para esse fim.

Portanto, ter a visão de quais pontos devem ser analisados e manter a técnica de análise em sigilo são duas facetas da ciência de dados que perpassam a economia digital. Há uma infinidade de dados pessoais que poderiam ser utilizados para inferir se determinada pessoa está grávida ou não, desde a compra de produtos para bebês até visitas em *sites* sobre gravidez e maternidade.

Essa limitação ocorre porque a vantagem competitiva das empresas emana justamente da capacidade que elas possuem de desenvolver essas metodologias de análise e produzir resultados mais acurados que os seus concorrentes.[393] Portanto, há um investimento significativo no desenvolvimento desses sistemas. Em paralelo, a abertura dessa informação pode inspirar os adversários, que podem se aproveitar dos investimentos do concorrente para aprimorar seus próprios serviços, sem ter os custos do aporte de capital. Portanto, por excelência, o segredo de negócio é um conhecimento que não se encontra no domínio público e que, a partir dele, a empresa que o obteve o mantém em confidencialidade e estabelece amarras contratuais com aqueles que partilham desse conhecimento[394]

A grande desvantagem do segredo de negócio é que nada impede que terceiros cheguem à mesma conclusão, através de suas próprias investigações. Por isso mesmo, o regime de proteção desses segredos é muito rígido, visto que está diretamente atrelado ao modelo de negócio das empresas de tecnologia. Além disso,

---

[392] HILL, K. How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did. *Forbes*, 16 fev. 2012. Disponível em: https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/. Acesso em: 27 ago. 2020.

[393] "Data on users is highly valuable for targeting digital advertising (particularly display advertising) and measuring its effectiveness. Advertisers and publishers have told us that Google and Facebook enjoy significant competitive advantages in both targeting and measuring effectiveness because of their extensive access to user data." (COMPETITION & MARKETS AUTHORITY. *Online platforms and digital advertising*: market study interim report. [*S. l.*]: CMA, 2019. Disponível em: https://assets.publishing.service.gov.uk/media/5dfa0580ed915d0933009761/Interim_report.pdf. Acesso em: 26 abr. 2021. p. 15).

[394] DUARTE, M. de F.; BRAGA, C. P. *Propriedade Intelectual*. 1. ed. [*S. l.*]: Sagah, 2018. p. 7.

amparado por tratados internacionais, as empresas podem fazer valer seu direito à confidencialidade tanto na esfera civil quanto penal[395]. A legislação de Propriedade Intelectual é altamente harmonizada ao redor do mundo, graças ao Acordo sobre Aspectos dos Direitos de Propriedade Intelectual Relacionados ao Comércio (TRIPS)[396], que, em seu art. 39, determina a proteção à informação confidencial como forma de combater a concorrência desleal.

O curioso é que a tutela legal do segredo industrial não decorre de uma previsão explícita da lei, mas de uma interpretação conjunta do art. 195, incisos XI e XII, da Lei de Propriedade Industrial brasileira[397], e do art. 39, § 2º, do TRIPs. Essa compreensão do segredo de negócio tem servido de instrumento para as empresas protegerem seus algoritmos nas mais diversas esferas. Fora do país, a agressividade das empresas em combater qualquer forma de intrusão é notória, sobretudo nos EUA. São essas proteções, que isolam os sujeitos das decisões automatizadas de qualquer acesso às suas "engrenagens" internas, que autores como Frank Pasquale[398], Burrel[399] e Rudin[400] chamam de "opacidade". É o conjunto de instrumentos organizacionais e jurídicos que tornam o complexo técnico das decisões efetivamente uma caixa preta (*black box*), e que é usado de forma ostensiva contra qualquer um que tenta descobrir um pouco mais sobre o funcionamento desses sistemas.

Outro ponto relevante ao se discutir as implicações dos direitos de propriedade intelectual sobre a proteção de dados diz respeito à natureza probabilística dos *outputs* de algoritmos. O ponto crucial de algumas estratégias de negócio, como a da Target e a identificação de clientes grávidas, está ligado a técnicas estatísticas. Algoritmos não são usados apenas para produzir resultados verificáveis sobre a vida dos titulares de direitos, mas também informações de natureza probabilística, que

---

395 COELHO, F. U. *Curso de Direito Comercial.* Direito de Empresa. São Paulo: Revista dos Tribunais, 2018. 2 v.

396 WORLD TRADE ORGANIZATION. T*rade-Related Aspects of Intellectual Property Rights*. 15. abr. 1994. Disponível em: https://wipolex.wipo.int/en/text/305907. Acesso em: 26 abr. 2021.

397 BRASIL. *Lei nº 9.279, de 14 de maio de 1976*. Regula direitos e obrigações relativos à propriedade industrial. Brasília, DF: Presidência da República, 14 maio 1996. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l9279.htm. Acesso em: 26 abr. 2021.

398 PASQUALE, F. *The Black box society:* the secret algorithms that control money and information. First Harvard University Press paperback edition ed. Cambridge, Massachusetts; London, England: Harvard University Press, 2015.

399 BURRELL, J. *How the Machine "Thinks:" Understanding Opacity in Machine Learning Algorithms*. Rochester, NY: Social Science Research Network, 2015. Disponível em: https://papers.ssrn.com/abstract=2660674. Acesso em: 27 maio. 2020.

400 RUDIN, C. Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead. *ArXiv [cs, stat]*, [*S. l.*], 2019. Disponível em: http://arxiv.org/abs/1811.10154. Acesso em: 10 ago. 2020.

orientam a empresa nas tomadas de decisão.

Por exemplo, um algoritmo de análise de crédito pode avaliar se o cliente “João” tem seu nome negativado em alguma base de dados. Esse dado produz uma constatação verificável: “João tem seu nome negativado em uma base de dados de proteção ao crédito.” Desse fato, a empresa pode avaliar se conceder crédito a esse indivíduo é algo que ela deve ou não fazer, assim como precificar essa concessão com base em uma leitura do risco que ela está assumindo. Por outro lado, essa constatação pode gerar inferências de natureza estatística, que orientam decisões às finalidades institucionais de quem implementa o algoritmo. Por exemplo, a inferência de que “João tem 70% de chance de inadimplir um empréstimo bancário”.

Essa segunda constatação é uma previsão acerca da vida do indivíduo, mas cuja realidade não é verificável[401]. É impossível medir se, de fato, um indivíduo tem 70% de chance de ser algo ou se encontrar em algum estado, pois isso é uma análise e não um fato. É possível afirmar que mesmo as avaliações verificáveis são, em alguma medida, sujeitas a erros e incertezas e por isso seriam também estatísticas. Ainda assim, há um limiar que se cruza quando estamos falando de chances e não de constatações que são fácil e imediatamente validadas.

Nesse sentido, as avaliações estatísticas são altamente voláteis, contingentes na série de fatores apreciados para se chegar àquele dado probabilístico. Nesse sentido, o dado pode ser confiável ou não confiável. Essa avaliação pode ser feita com base na metodologia estatística e nos dados que foram usados para chegar a esse resultado.

Por exemplo, é possível afirmar que Maria tem 50% de chance de vestir vermelho e 50% de vestir azul na segunda-feira, se partirmos de algumas premissas importantes. Primeiro, a de que a Maria só usará um desses dois trajes; segundo, de que ela escolhe a roupa jogando uma moeda, onde cada lado corresponde a uma das decisões; terceiro, de que essa moeda tem de fato 50% de chance de cair em cada um dos lados. Portanto, a conclusão estatística de que Maria tem 50% de chance de usar vermelho na segunda-feira não é uma realidade, mas é uma avaliação de possibilidades que é diretamente dependente da qualidade e veracidade das premissas que foram usadas nessa avaliação. A avaliação só faz sentido se tivermos

[401] WACHTER, S.; MITTELSTADT, B. A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI. *LawArXiv*, 12 out. 2018. Disponível em: https://doi.org/10.31228/osf.io/mu2kf. Acesso em: 27 maio 2020.

também ciência de como esse número é construído.

A proteção jurídica dos algoritmos como segredo comercial é, então, muito atrelada a essa premissa implícita de que dados estatísticos são avaliações subjetivas que as empresas fazem para sua própria estratégia de negócio. Contudo, essa percepção de que o dado estatístico se resume a uma questão de avaliação subjetiva de estratégia comercial é equivocada. No contexto de tratamentos automatizados, essas inferências estatísticas desencadeiam ou orientam decisões que afetam diretamente direitos individuais. Os exemplos desses problemas são inúmeros, incluindo o mercado de crédito[402], a concessão de vistos para a entrada de estrangeiros[403] e mesmo análises de CVs de candidatos[404]. Essas avaliações estatísticas, quando acopladas a estruturas de decisão, podem desencadear políticas discriminatórias ilegais de toda ordem.

No caso de decisões automatizadas com base em dados não estatísticos, a remediação dessas dificuldades é dada de forma mais objetiva. Se um direito é assegurado com base em uma informação verificável, então é simples detectar um erro na decisão. Para assegurar a qualidade das decisões, surge o direito do titular de corrigir os dados: trata-se de uma proteção sobre a veracidade das informações (art. 18, III, da LGPD).

Essa é uma distinção fundamental dada por Wachter & Mittelstadt[405]: a definição de dado pessoal abarca o conceito de inferências estatísticas, mas o direito de correção se vê limitado pela verificabilidade das informações registradas. Assim, um dado estatístico, por mais que fosse referente a pessoa identificada ou identificável, se tornaria uma subcategoria de dados pessoais, onde o titular gozaria de menos prerrogativas de proteção e controle.

A confiabilidade da análise estatística é muito atrelada aos processos de análise empregados, e por isso são majoritariamente compreendidos não como um

---

[402] KLEIN, A. Reducing bias in AI-based financial services. *BROOKINGS*. 10 jul. 2020. Disponível em: https://www.brookings.edu/research/reducing-bias-in-ai-based-financial-services/. Acesso em: 1 set. 2020.

[403] HOME Office drops “Racist” algorithm from Visa Decisions. *BBC News*, [*S. l.*], 4 ago. 2020. Disponível em: https://www.bbc.com/news/technology-53650758. Acesso em: 27 ago. 2020.

[404] MARTINEZ, A. Considering AI In Hiring? As Its Use Grows, So Do The Legal Implications For Employers. *Forbes*, 5 dez. 2020. Disponível em: https://www.forbes.com/sites/alonzomartinez/2019/12/05/considering-ai-in-hiring-as-its-use-grows-so-do-the-legal-implications-for-employers/#45c4dca77d47. Acesso em: 1 set. 2020.

[405] WACHTER, S.; MITTELSTADT, B. A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI. *LawArXiv*, 12 out. 2018. Disponível em: https://doi.org/10.31228/osf.io/mu2kf. Acesso em: 27 maio 2020.

direito objetivo do titular, mas segredo de negócio. Esse regime de confidencialidade, protegido por instrumentos jurídicos privados, coloca uma grande barreira à capacidade dos titulares de dados de fazerem valer seus direitos. Na Europa, como mostram Wachter e Mittelstadt[406], há intenso debate sobre se essas inferências teriam qualquer proteção na esfera de proteção de dados pessoais, e as cortes europeias tendem, por enquanto, a acreditar que não há.

Na Europa há, de um lado, o entendimento do Article 29 Working Party que interpreta o arts. 4º e 9º da GDPR como substrato jurídico suficiente para estabelecer que inferências podem ser dados pessoais, tendo em vista que o texto legal define dados pessoais como informação relativa à pessoa identificada ou identificável. Ainda que essas informações não sejam "fornecidas" pelos usuários, elas são derivadas ou inferidas a partir de outras informações. Assim sendo, propõe-se um teste de três elementos para se avaliar se dados não pessoais podem ser considerados pessoais: o conteúdo, propósito ou resultado do processamento de dados precisa ter alguma relação com o indivíduo.

O último elemento do teste, "resultado", é um elemento chave na definição jurídica do valor das inferências. O critério que se extrai desse elemento é que qualquer resultado que impacte a esfera de interesses do indivíduo é passível de ser considerado dado pessoal, pela definição legal e pela própria função que a legislação concebe no instrumento de proteção de dados. A dúvida jaz na avaliação de inferências como "informações, opiniões ou avaliações" subjetivas e não verificáveis. Esses critérios são o ponto de conflito maior, pois o Article 29 Working Party, junto com alguns juristas, acredita nesse alargamento conceitual[407].

Nesse campo, a Corte Europeia de Justiça (ECJ) apresenta uma visão diferente sobre o tema, estabelecendo, ao menos por enquanto, uma visão muito mais restritiva para o conceito de dados pessoais. A ECJ, nos casos C-141/12 e C-372/12, avaliou o direito do titular de dados de ter informação sobre a avaliação feita por um sistema que determinava a concessão de residência em um país. A Corte e o Advogado Geral

[406] WACHTER, S.; MITTELSTADT, B. A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI. *LawArXiv*, 12 out. 2018. Disponível em: https://doi.org/10.31228/osf.io/mu2kf. Acesso em: 27 maio 2020.

[407] WACHTER, S.; MITTELSTADT, B. A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI. *LawArXiv*, 12 out. 2018. Disponível em: https://doi.org/10.31228/osf.io/mu2kf. Acesso em: 27 maio 2020.

se manifestaram no sentido de restringir a proteção de dados pessoais[408].Eles argumentaram que o escopo da legislação de proteção de dados não poderia ser alargado a tal ponto. Primeiro, por conta da própria definição de dados pessoais, que eles entendem ser mais restrita. Segundo, das limitações oferecidas pelo direito de acesso e retificação da informação. Por fim, e mais importante, a visão geral de que o regulamento europeu não serve para assegurar que o processo decisório automatizado seja preciso ou lícito, não tendo, portanto, o condão de governar inferências.

No Brasil, a indefinição é ainda maior, visto que a entrada em vigor da LGPD é ainda mais recente e falta jurisprudência e manifestações de entidades reguladoras para dirimir essas dúvidas no país. É por conta de demandas semelhantes que a própria legislação consumerista brasileira, assim como os julgados do STJ abordados anteriormente nesta tese, estabelecem direitos para os titulares compreenderem os critérios que dão lugar às decisões automatizadas. No mínimo, compreende-se quais são os elementos que são levados em conta, mas isso não explica exatamente o procedimento e o conjunto de decisões que faz avaliações por excelência comparativas e discriminatórias a partir dos dados coletados.

Como veremos no julgado envolvendo o MPF-RJ e a empresa Decolar, na seção 4.3.3.1, a defesa do instituto do segredo industrial é bastante prezada pelas empresas de tecnologia, sobretudo porque é algo que viabiliza seu atual modelo de negócios. Esse fato também se evidencia na própria LGPD, que foi alvo de enorme pressão pelo setor industrial para incluir diversas proteções ao segredo comercial no seu texto.[409] Nesse sentido, uma simples comparação da proteção desse instituto no regulamento europeu e na lei brasileira dá uma dimensão do relevo que ele assume no cenário brasileiro. Na GDPR, há apenas uma menção ao segredo comercial ou propriedade intelectual, no Considerando 63[410], no que tange ao direito de acesso do

---

[408] WACHTER, S.; MITTELSTADT, B. A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI. *LawArXiv*, 12 out. 2018. Disponível em: https://doi.org/10.31228/osf.io/mu2kf. Acesso em: 27 maio 2020.

[409] INTERNETLAB. O que está em jogo no debate sobre dados pessoais no Brasil? Relatório final sobre o debate público promovido pelo Ministério da Justiça sobre o Anteprojeto de Lei de Proteção de Dados Pessoais. São Paulo: InternetLab, 2016. Disponível em: https://www.internetlab.org.br/wp-content/uploads/2016/05/reporta_apl_dados_pessoais_final.pdf. Acesso em: 4 maio 2021.

[410] UNIÃO EUROPEIA. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance)*. Official Journal of the European Union, 4 maio 2016. Disponível em; https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng. Acesso em: 27 ago. 2021.

titular dos dados. No contexto europeu, o direito do titular de acessar os dados coletados sobre ele, assim como as formas de processamento que vêm sendo feitas com aqueles dados, é garantido por lei. O Considerando 63 apenas assegura que esse acesso não viole outras liberdades, incluindo o segredo comercial e o direito autoral, mas que isso não isenta o processador de fornecer as informações demandadas pelo titular.

No Brasil, a LGPD, normativa de extensão muito menor quando comparada à GDPR, o termo segredo comercial aparece em 14 passagens, na maioria das vezes também como ressalva ao acesso às informações sobre processos de tratamento de dados pessoais. Esse número de recorrências demonstra a preocupação que o legislador brasileiro teve em proteger o segredo comercial, nitidamente muito maior do que o legislador europeu. Essa decisão do legislador pode ser em grande medida atribuída à movimentação do setor empresarial no processo legislativo da LGPD, que buscou resguardar seus interesses no contexto de uma nova norma abrangente sobre proteção de dados pessoais.[411]

4.1.2 Limitações do conceito de dados pessoais

O art. 5º, inciso I, da LGPD, estabelece que dado pessoal é uma informação relacionada a uma pessoa natural identificada ou identificável. No limite, qualquer informação que possa ser atribuída a um indivíduo tem enquadramento na norma, ainda que a pessoa não tenha fornecido esse dado diretamente ao processador. O enquadramento de dado pessoal advém de um fato jurídico, ou seja, a existência dessa informação sobre pessoa identificada ou identificável, e não de um ato de coleta.

Essa distinção mostra que a definição de dado pessoal, que foi positivada pela LGPD, baseia-se na concepção expansionista de dado pessoal. O viés reducionista é a dimensão mais precisa do dado, que versa sobre pessoa específica, determinada, com vínculo direto à informação[412]. Contudo, o uso de termo “identificável” insere um grau de indeterminação que alarga o conceito de dado pessoal, oferecendo também

[411] INTERNETLAB. O que está em jogo no debate sobre dados pessoais no Brasil? Relatório final sobre o debate público promovido pelo Ministério da Justiça sobre o Anteprojeto de Lei de Proteção de Dados Pessoais. São Paulo: InternetLab, 2016. Disponível em: https://www.internetlab.org.br/wp-content/uploads/2016/05/reporta_apl_dados_pessoais_final.pdf. Acesso em: 4 maio 2021.
[412] BIONI, Bruno Ricardo, 2019.

um direito a pessoas indeterminadas, sem vínculo direto com a informação registrada[413].

Seguindo o texto da lei, mesmo dados estatísticos sobre uma pessoa podem ser considerados dados pessoais. Pelo critério, não é a verificabilidade do dado que determina se é ou não pessoal, mas o fato dele guardar relação com uma pessoa identificável. A lei incide sobre um fato jurídico – a existência do dado relativo a uma pessoa — e não sobre um ato específico que seja feito com os dados.

Na Europa há, de um lado, um entendimento de que seja possível[414], a partir de uma análise sistemática dos documentos do Article 29 Working Party, posteriormente adotados pelo European Data Protection Board[415], dos arts. 4º e 9º da GDPR, estabelecer que inferências podem ser dados pessoais, tendo em vista que o texto legal define dados pessoais como informação relativa à pessoa identificada ou identificável. Ainda que essas informações não sejam "fornecidas" pelos usuários, elas são derivadas ou inferidas a partir de outras informações. Assim sendo, propõe-se um teste de três elementos para se avaliar se dados não pessoais podem ser considerados pessoais: o conteúdo, propósito ou resultado do processamento de dados precisa ter alguma relação com o indivíduo.

O último elemento do teste, "resultado", é um elemento chave na definição jurídica do valor das inferências. O critério que se extrai desse elemento é que qualquer resultado que impacte a esfera de interesses do indivíduo é passível de ser considerado dado pessoal, pela definição legal e pela própria função que a legislação concebe no instrumento de proteção de dados. A dúvida jaz na avaliação de

---

413 BIONI, B. Xeque-Mate: o tripé de proteção de dados pessoais no xadrez das iniciativas legislativas no Brasil. *GPOPAI/USP*, [*S. l.*], 2015. Disponível em: https://www.researchgate.net/publication/328266374_Xeque-Mate_o_tripe_de_protecao_de_dados_pessoais_no_xadrez_das_iniciativas_legislativas_no_Brasil. Acesso em: 27 ago. 2020.

414 WACHTER, S.; MITTELSTADT, B. A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI. *LawArXiv*, 12 out. 2018. Disponível em: https://doi.org/10.31228/osf.io/mu2kf. Acesso em: 27 maio 2020.

415 EUROPEAN DATA PROTECTION BOARD. *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 679/2016*. Bruxelas: European Commission, 2016. Disponível em: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053; EUROPEAN DATA PROTECTION BOARD. *Guidelines 2/2019 on the processing of personal data under Article 6(1)(b) GDPR in the context of the provision of online services to data subjects*. [*S. l.*]: EDPB, 2019. Disponível em: https://privacyblogfullservice.huntonwilliamsblogs.com/wp-content/uploads/sites/28/2019/04/edpb_draft_guidelines-art_6-1-b-final_public_consultation_version_en.pdf. Acesso em: 26 abr. 2021; e EUROPEAN DATA PROTECTION BOARD. *ARTICLE 29 Newsroom — Guidelines on the right to "data portability"*. Disponível em: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611233. Acesso em: 26 abr. 2021.

inferências como “informações, opiniões ou avaliações” subjetivas e não verificáveis. Esses critérios são o ponto de conflito maior, pois o Article 29 Working Party, junto com alguns juristas, acreditam nesse alargamento conceitual[416].

No entanto, alguns julgados anteriores apresentam visões mais restritas. Conforme mencionada no subcapítulo anterior, a Corte Europeia de Justiça, nos casos C-142/12 e C-372/12 sobre a concessão de residências, entendeu que a proteção de dados pessoais não poderia interferir na correção das inferências. Tal interpretação, segundo Wachter, representa um enfraquecimento da proteção dos direitos dos titulares de dados no contexto de *big data*[417].

O enfraquecimento da proteção de dados se agrava quando tratamos não de indivíduos, mas de grupos. Por carregar inferências relacionais, é perfeitamente possível produzir decisões que discriminem grupos. Essa é a lógica exata de serviços de personalização fundados em uma base comparativa como “usuários que gostaram de produto X também gostarão de produto Y”. O agrupamento de usuários pode ocorrer sem passar diretamente por dados pessoais, e na verdade os danos das decisões automatizadas pode ocorrer pela formação de grupos e assim contornar o regime individualizado da proteção de dados[418]. Por isso, a medida relacional derivada entre os dados e, portanto, entre os indivíduos, precisa ser uma proteção essencial do indivíduo, o que pode ser garantido por meio da dimensão consequencialista do dado pessoal[419], ainda que não haja sua previsão explícita no texto legal.

Percebemos que há uma certa recursividade no campo de sistemas automatizados. É impossível compreender um dado sem se entender como ele foi obtido, então o embate travado no seio da própria LGPD aponta para uma antinomia. A proteção de dados não consegue ser completa, dentro do seu escopo expansionista, sem que se imponha obrigações de transparência aos algoritmos, visto

---

[416] EUROPEAN DATA PROTECTION BOARD. *ARTICLE 29 Newsroom — Guidelines on the right to “data portability”*. Disponível em: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611233. Acesso em: 26 abr. 2021.

[417] WACHTER, S.; MITTELSTADT, B. A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI. *LawArXiv*, 12 out. 2018. Disponível em: https://doi.org/10.31228/osf.io/mu2kf. Acesso em: 27 maio 2020.

[418] FLORIDI, L. Open Data, Data Protection, and Group Privacy. *Philosophy & Technologys.* v. 27, n. 1, p. 1–3, 2014. Disponível em: https://doi.org/10.1007/s13347-014-0157-8. Acesso em: 21 jun. 2021.

[419] BIONI, B. Xeque-Mate: o tripé de proteção de dados pessoais no xadrez das iniciativas legislativas no Brasil. *GPOPAI/USP*, [*S. l.*], 2015. Disponível em: https://www.researchgate.net/publication/328266374_Xeque-Mate_o_tripe_de_protecao_de_dados_pessoais_no_xadrez_das_iniciativas_legislativas_no_Brasil. Acesso em: 27 ago. 2020.

que o conteúdo relacional dos dados é resultado dos procedimentos aos quais eles são submetidos.

Neste sentido, podemos observar uma diferença significativa entre a GDPR e a LGPD. No regulamento europeu, por exemplo, os dados anonimizados não são protegidos, estão fora do escopo de aplicação da norma. O Recital 26 estabelece um conceito de anonimização e ressalta que processos passíveis de reversão deverão ser considerados como pseudoanonimização e, portanto, dados pessoais. Neste sentido, permite que determinadas categorias de dados sejam excluídas da proteção de dados pessoais. Há objeções importantes sobre a possibilidade de uma anonimização efetiva. Além disso, essa categorização exclui dados anonimizados que eventualmente sejam utilizados para realizar inferências estatísticas sobre determinados grupos sociais. O art. 12 da LGPD, por sua vez, no § 2°, estabelece que mesmo dados anonimizados poderão ser considerados dados pessoais quando forem utilizados para a construção de perfis que afetem pessoas naturais. Neste sentido, em vista das consequências advindas do tratamento no livre desenvolvimento da personalidade de pessoas naturais, os dados utilizados em algoritmos, mesmo que por si só não sejam caracterizados como pessoais, poderão ser considerados pessoais, mesmo que sejam anonimizados.

Essa possível interpretação do art. 12 permite afirmar que a LGPD teria condições de tutelar efeitos não previstos no GDPR, incluindo a razoabilidade das inferências realizadas sobre pessoas naturais. No entanto, por mais que a interpretação em favor da proteção mais ampla de dados pessoais seja robusta, ainda há uma intensa disputa sobre se inferências estatísticas configuram dados pessoais. Esse alargamento ainda não tem definição clara no ordenamento brasileiro, mas carrega decorrências importantes no escopo do direito à explicação.

Primeiro, porque o direito à explicação passa a valer para um titular indefinido e, eventualmente, difuso ou coletivo. Portanto, existiria não só um direito subjetivo individual de demandar os critérios de decisões automatizadas, mas um direito difuso, que pode ser usado em nome de interesses que não são individuais ou individualizáveis. Passa a existir um interesse mais amplo no funcionamento do processo decisório, e não apenas no caso concreto onde decisões podem ter um efeito lesivo.

A segunda consequência, que é um resultado da primeira, é que quanto mais se expandir as garantias dos titulares de dados pessoais sobre as relações e

inferências que surgem dos seus dados, cada vez mais se invadirá o espaço que é hoje dominado pelo segredo comercial. Esse espaço, como vemos pela própria movimentação no processo legislativo da LGPD, é um ponto onde as empresas estão dispostas a proteger com afinco a sua competitividade estratégica. Portanto, o conceito de dado pessoal reconhecido pela legislação e a sua interpretação são essenciais para definir em quais contextos o direito à explicação será garantido. Se se defende que este deve ser garantido quando da existência de processos automatizados baseados no tratamento de dados pessoais, a *contrario sensu*, ele não existiria em procedimentos automatizados baseados em dados não-pessoais. O conceito, portanto, limita o escopo de aplicação da norma e, por consequência, o próprio direito à explicação.

4.2 DESAFIOS COGNITIVOS: LIMITAÇÕES HUMANAS À COMPREENSÃO DE SISTEMAS COMPLEXOS

Computadores possuem uma característica enigmática. Essa característica foi apontada por Bennet[420] como um dos fatores que explicam a convergência de regulação em diferentes lugares. São caixas pretas de difícil compreensão para pessoas comuns. Um dos fatores comuns na motivação para regular os dados seria o desejo dos legisladores de tomarem o controle do processamento de informações pessoais nos meios digitais.

O conhecimento sobre os sistemas de informação e os meios computacionais pode ocorrer em diversas camadas e de formas diferentes. Pode haver recortes geracionais ou sociais em decorrência do nível de conhecimento da população. A capacidade de compreensão do universo digital cumpre um papel importante no exercício do direito à privacidade e na sua regulação.

Se na opacidade derivada de uma estratégia corporativa o direito à explicação encontra limites relacionados ao acesso aos códigos dos algoritmos. Por vezes, mesmo com o conhecimento desses códigos, há limitações ao direito à explicação derivadas de questões cognitivas dos sujeitos que deveriam analisar tais códigos. Jenna Burrell[421] aponta como a natureza técnica e especializada da computação torna

[420] BENNET, C. *Regulating privacy*: data protection and public policy in Europe and the United States. Ithaca, New York: Cornell University Press, 1992.
[421] BURRELL, J. *How the Machine "Thinks:" Understanding Opacity in Machine Learning Algorithms*. Rochester, NY: Social Science Research Network, 2015. Disponível em: https://papers.ssrn.com/abstract=2660674. Acesso em: 27 maio. 2020.

a compreensão dos algoritmos difícil para não especialistas e, além disso, como alguns modelos de aprendizado de máquina (*machine learning*) são opacos até mesmo para quem os desenvolve.

Dessa forma, a aprovação do GDPR levantou questões importantes para a comunidade das ciências computacionais. Essa seção pretende discutir as especificidades das tecnologias. Em primeiro lugar é preciso discutir a natureza especializada das tecnologias de informação e das linguagens computacionais, as maneiras pelas quais podem ser conhecidas, bem como os limites do conhecimento para não especialistas.

Então podemos discutir de forma específica os limites da explicação no campo do *machine learning*. Nos últimos anos, uma série de trabalhos na área da ciência da computação vem discutindo o problema da aplicabilidade de modelos algorítmicos complexos e propondo metodologias para o desenvolvimento de *explainable artificial intelligence* (XAI).

Por fim, discutiremos como o direito à explicação se apresenta frente ao conhecimento geral da população sobre computação, matemática, ciência etc. Discutiremos alternativas para o acesso à explicação para a heterogeneidade dos grupos sociais. Devemos argumentar como o comprometimento das organizações é essencial para o exercício dos direitos no contexto da desigualdade de conhecimento sobre tecnologias da informação, conhecida como *digital divide*.

4.2.1 A complexidade de sistemas algorítmicos

É comum afirmar que as ciências explicam o funcionamento do universo. Os modelos científicos são as formas mais comuns que os cientistas utilizam para explicar a realidade. De forma muito simplificada, podemos afirmar que os modelos são representações que encadeiam determinadas proposições de forma estruturada para representar uma certa realidade[422].

No entanto, essa explicação nunca será tão complexa como a própria realidade. Um conto de Borges[423], muito lembrado, narra a história de um império que

[422] STREVENS, M. No understanding without explanation. *Studies in History and Philosophy of Science Part A*. v. 44, n. 3, p. 510–515, 2013. Disponível em: https://doi.org/10.1016/j.shpsa.2012.12.005. Acesso em: 21 jun. 2021.
[423] BORGES, J. L. Sobre o Rigor na Ciência. *In*: BORGES, J. L. *História Universal da Infâmia*. Lisboa: Assírio e Alvim, 1982. p. 117.

se dedicava à cartografia com muito afinco, desenhando mapas mais próximos possíveis da realidade. Dessa forma, os mapas de uma província eram do tamanho de uma cidade, enquanto os mapas do império, do tamanho de uma província. Os cartógrafos chegaram ao ponto de criar um mapa do império tão perfeito que tinha exatamente o mesmo tamanho do império, coincidindo com ele ponto por ponto. Com o passar do tempo, as gerações posteriores concluíram que tais representações eram inúteis e os mapas transformaram-se em ruínas que passaram a abrigar mendigos e animais.

Nas ciências os modelos são utilizados para explicar fenômenos e as relações entre eles. São simplificações, de forma que nenhum deles é completo. Referem-se a domínios específicos da realidade. A física newtoniana, por exemplo, funciona para corpos com dimensões e massa significativas em contextos de velocidades que não se aproximam da velocidade da luz. Os modelos atômicos são importantes para explicar o comportamento de algumas partículas, mas não explicam outros fenômenos subatômicos. Os modelos econômicos de oferta e demanda são ótimos para descrever o funcionamento de preços em um mercado competitivo e sem variações qualitativas, mas nem sempre funcionam nos mercados efetivos, em transformação, regulados ou concentrados[424].

Nesse sentido, é comum afirmar que as explicações através dos modelos devem ser corretas, não necessariamente verdadeiras. Os cientistas algumas vezes necessitam simplificar e idealizar modelos que não correspondem estritamente à realidade. Na química pode-se desprezar as forças intermoleculares, nos estudos em biologia afirma-se que as populações são infinitas em modelos evolutivos. Na economia afirma-se que os seres humanos tomam ações racionais. Embora não correspondam à realidade, argumenta-se que tais proposições e simplificações apenas significam que tais elementos não fazem diferença para determinado problema específico[425].

Essas construções carregam em si uma complexidade que só pode ser efetivamente compreendida após determinado esforço e engajamento do sujeito em

[424] STREVENS, M. No understanding without explanation. *Studies in History and Philosophy of Science Part A*. v. 44, n. 3, p. 510–515, 2013. Disponível em: https://doi.org/10.1016/j.shpsa.2012.12.005. Acesso em: 21 jun. 2021.
[425] STREVENS, M. No understanding without explanation. *Studies in History and Philosophy of Science Part A*. v. 44, n. 3, p. 510–515, 2013. Disponível em: https://doi.org/10.1016/j.shpsa.2012.12.005. Acesso em: 21 jun. 2021.

torno de determinado campo científico. A compreensão desses fenômenos depende do conhecimento dos principais problemas, do contexto da criação de teorias, modelos e leis, bem como a prática de reflexão entre os pares, incluindo o conhecimento das limitações, contradições e debates.

Esse raciocínio se aplica também aos algoritmos. Programadores constroem códigos baseados em diversas teorias e ramos do conhecimento científico, em linguagem lógica e matemática, de forma que a mera divulgação das linhas e comandos não permite a sua compreensão. Para cada operação inscrita no código há premissas implícitas, teorias e pressuposições, bem como objetivos. Há, além disso, escolhas do programador do algoritmo sobre quais dados ou informações utilizar. Nesse sentido, para a compreensão do algoritmo é preciso compreender como opera um computador e as regras e leis de um campo específico do conhecimento do qual o algoritmo se apropria para realizar atividades de forma automatizada.

4.2.1.1 Opacidade inerente aos modelos complexos

Anthony Giddens apresenta em seu livro uma importante constatação sobre a modernidade. Segundo o sociólogo, a modernidade traz mudanças institucionais e, acima de tudo, na forma como a experiência cotidiana é representada e vivida pelos sujeitos. A modernidade é uma ordem pós-tradicional, onde as relações calcadas no conhecimento compartilhado e reproduzido de geração em geração dá lugar a uma profunda racionalização dos mais amplos aspectos da vida mediados por novas tecnologias e arranjos sociais diversos[426].

Conforme as relações sociais se espraiam no tempo e no espaço a partir do processo de globalização, as fontes de autoridade se multiplicam. Conforme criam-se novos corpos de conhecimento acumulado, reduz-se a segurança ontológica e criam-se constantes possibilidades de fragmentação dos sujeitos num regime de dúvida constante. A sociedade do risco emerge não porque a vida seja mais perigosa, mas porque cada vez mais as decisões necessárias demandam cálculos de risco. E esses cálculos, contudo, não podem se realizar a partir unicamente da experiência individual.

Nesse sentido, a vida moderna é cada vez mais permeada pela atuação de especialistas, de corpos técnicos e de saberes científicos. É praticamente impossível

[426] GIDDENS, A. *Modernidade e identidade*. Rio de Janeiro: Zahar, 2002.

tomar as decisões importantes na sociedade moderna sem o auxílio direto de especialistas. Cria-se um interessante paradoxo no qual é cada vez mais fácil aprender e tornar-se especialista em um tema, e, ao mesmo tempo, cada vez mais difícil compreender mais do que algumas especialidades. Dessa forma, cada campo do conhecimento torna-se opaco para a grande maioria da população não especialista de forma que há uma opacidade inerente em qualquer modelo mais especializado ou complexo em todos os campos da ciência.

As tecnologias da informação têm um caráter generalista e encontram aplicações nos mais diversos campos da ciência. As teorias e modelos científicos de cada campo do conhecimento encontram dados e capacidade computacional para realização de tarefas cada vez mais complexas do nosso dia a dia. Dessa forma, somos permeados por uma série de assunções, modelos e pressupostos científicos dos quais muitas vezes sequer temos conhecimento.

Alguns modelos estabelecem relações causais entre fenômenos, leis ou regras gerais a partir da observação da realidade. Em alguns casos pode ser muito simples descobrir a relação entre dois eventos, por exemplo, entre soltar um objeto de determinada altura e ter como consequência a sua queda. Em outros casos essa relação é mais difícil de se estabelecer. Algumas ciências observam fenômenos que podem ter mais de um causa, ou que apresentam, sob as mesmas condições iniciais, resultados completamente diferentes. Nessas ocasiões as ciências costumam estabelecer relações entre fenômenos a partir de cálculos estatísticos.

A estatística é um ramo da matemática para a descrição e interpretação de massas de informações numéricas. Pode-se analisar um número de casos de determinados eventos e estabelecer proporções e probabilidades. Nas condições em que não é possível estabelecer uma relação de efeitos necessários a partir de uma causa pela diversidade de fenômenos, pode-se calcular a probabilidade de dois eventos ocorrerem ao mesmo tempo observando um grande volume de eventos, de forma que se possa conhecer o que são os fenômenos mais ou menos comuns e o que são exceções[427].

Uma das medidas mais comuns em estatística é a que se pode estabelecer entre eventos ou variáveis, a que se chama correlação. Numa escola, por exemplo, pode-se observar as horas de estudo de cada aluno com as notas no exame de

[427] DEVORE, J. L. *Probabilidade e estatística*: para engenharia e ciências. Tradução: Joaquim Pinheiro Nunes da Silva. São Paulo: Cengage Learning, 2006. p. 45–75.

matemática. Uma delas, as horas de estudo, chamamos variável independente, outra, as notas, chamaremos de variável dependente. Alguns alunos estudam mais, outros menos. Cada aluno apresenta uma nota diferente. Alguns dos alunos podem estudar poucas horas e terem notas altas, outros, estudarem por muitas horas e terem uma nota menor. No entanto, ao se observar um número grande de alunos, poderemos estabelecer essa correlação entre essas duas variáveis. Se a nota tende a crescer quando observamos mais horas de estudos, podemos dizer que há uma correlação positiva entre essas duas variáveis[428].

O uso de computadores contribuiu muito para a evolução da estatística. Pela capacidade de realizar uma grande quantidade de operações lógicas, os computadores permitem processar uma quantidade enorme de dados em pouco tempo. Basta imaginar a dificuldade de se calcular a proporção de homens e mulheres em uma população de um país a partir de fichas de papel. Os computadores modernos são capazes de realizar essa atividade em frações de segundo.

Além disso, outro dado importante sobre os cálculos estatísticos é que possuem um grau de precisão que aumenta conforme mais dados estão disponíveis. Esse é outro motivo do porquê o campo da estatística evolui a partir da computação. A emergência da sociedade da informação trouxe a multiplicação de dados dos mais diferentes tipos e naturezas, o que possibilitou a construção de modelos para explicar uma série de fenômenos. Recentemente essas atividades emergiram como um campo autônomo na ciência de dados, com aplicações em vários campos do conhecimento ou setores. O desenvolvimento de muitas ciências consiste na discussão desses números e correlações, confrontados com teorias e hipóteses, entre agentes que conhecem e compreendem esses problemas e suas limitações. Os modelos então são apresentados, discutidos, aceitos, superados ou atualizados no debate dentro de um campo científico.

O tipo de conclusões baseadas em análises estatísticas e probabilidades têm uma série de aplicações nos setores econômicos. Na análise de crédito, por exemplo, diante da vasta disponibilidade de informações financeiras, pode-se calcular a probabilidade de pagamento ou inadimplemento de empréstimos a partir de determinadas variáveis sobre os solicitantes. Nos planos de saúde, com informações relacionadas a hábitos de consumo ou estilos de vida, pode-se relacionar a maior ou

[428] DEVORE, J. L. *Probabilidade e estatística*: para engenharia e ciências. Tradução: Joaquim Pinheiro Nunes da Silva. São Paulo: Cengage Learning, 2006. p. 432–461.

menor probabilidade de manifestação de uma doença a partir de determinados comportamentos. As mesmas informações, quando inseridas em algoritmos, podem inclusive ser utilizadas para orientar decisões automatizadas.

Se os modelos não são novidade no campo da ciência e nos meios técnicos, nas últimas décadas o debate tem se tornado mais presente em função de aplicações comerciais em áreas que podem afetar direitos e liberdades[429]. Se o debate entre a precisão ou adequação de modelos dentro da academia se faz por um processo transparente e aberto, em contextos comerciais esses modelos se destinam ao uso por pessoas que não são cientistas ou técnicos, e, além disso, alguns estabelecem decisões não passíveis de questionamento em seus fundamentos, visto que o modelo que as embasa ser de natureza privada[430].

Os algoritmos realizam classificações, inferências e mesmo decisões e seu processo não é necessariamente explicado. Se nas relações jurídicas (legais ou contratuais) entre seres humanos há mecanismos e um corpo doutrinário que permite questionar as decisões e responsabilizar os agentes, esse processo ainda não se desenvolveu completamente para decisões automatizadas feitas por algoritmos.

Os algoritmos compõem uma forma de modelo semelhante aos utilizados nas ciências[431]. Nesse sentido, sua compreensão depende de conhecimentos técnicos e um engajamento do sujeito dentro de determinado campo científico e isso é um desafio ao direito à explicação. Trataremos no item 4.2.2 sobre as formas de traduzir modelos complexos para uma linguagem acessível a pessoas comuns, mas, antes disso, devemos tratar de aplicações específicas de algoritmos que possuem uma opacidade até mesmo para os especialistas.

4.2.1.2 IA, ML e interpretabilidade

Os modelos de inteligência artificial são sistemas capazes de realizar atividades para as quais um ser humano precisaria refletir, ou seja, utilizar a

---

[429] O'NEIL, C. *Weapons of Math Destruction:* How Big Data Increases Inequality and Threatens Democracy. New York, NY: Crown Books, 2016. *E-book*.

[430] DIAKOPOULOS, N. *Algorithmic accountability reporting:* on the investigation of black boxes. [*S. l.*]: Tow Center for Digital Journalism, 2014. Disponível em: https://academiccommons.columbia.edu/doi/10.7916/D8ZK5TW2. Acesso em: 21 jun. 2021.

[431] MITTELSTADT, B.; RUSSELL, C.; WACHTER, S. Explaining Explanations in AI. *In*: THE CONFERENCE, 2019, Atlanta, GA, USA. *Proceedings of the Conference on Fairness, Accountability, and Transparency — FAT'19*. Atlanta, GA, USA: ACM Press, 2019. p. 279–288. Disponível em: https://doi.org/10.1145/3287560.3287574. Acesso em: 27 maio 2020.

inteligência. No entanto se mimetizam atributos da inteligência humana, reagindo a estímulos exteriores de forma a resolver problemas, estes sistemas não elaboram esses processos de uma forma humanamente compreensível. Todavia, há algoritmos no campo da inteligência artificial cujos códigos são compostos por processos complexos, oriundos de um número gigantesco de processos automáticos e, portanto, não são facilmente interpretáveis[432].

A ideia de desenvolvimento de uma inteligência artificial remonta à antiguidade e sempre esteve presente na literatura, nas artes e na filosofia, desde Homero, passando por Descartes, Leibniz, Julio Verne e Isaac Asimov[433]. No entanto, o surgimento da computação e da robótica estabeleceu um horizonte muito promissor e mais concreto para aquelas fantasias. A possibilidade de realizar uma série de operações lógicas levou os cientistas a indagar o que seria necessário para fazer o computador pensar.

Os primeiros modelos de inteligência artificial eram baseados em representações simbólicas de forma a se assemelhar aos processos cognitivos e racionais. Os modelos tentavam codificar certo tipo de conhecimento de forma que possa ser aplicado, por um modelo, a partir de uma linguagem de computação, na resolução de problemas. Desde os anos 50 até os anos 80, essa foi a abordagem dominante em aplicações de busca, descoberta e planejamento, principalmente em aplicações especializadas em algumas áreas da ciência. Sua alta dependência de proposições hierarquicamente estruturadas e da precisão nos *inputs* tornam sua aplicação prática inviável na realidade de informações incompletas e imprecisas.

Alguns desses modelos podem ser interpretados visto ser possível observar as etapas do modelo criado pelo treinamento ou mesmo a partir das variáveis de entrada e de saída. Outros modelos não podem ser facilmente compreendidos, pois consistem em camadas de processos compostas por operações computacionais não interpretáveis por seres humanos[434]. Os algoritmos de visão computacional, por exemplo, fornecem um exemplo desse tipo. A partir de uma certa amostra de imagens de determinado objeto, o algoritmo produz um modelo capaz de identificar e classificar

---

[432] BURRELL, J. *How the Machine "Thinks:" Understanding Opacity in Machine Learning Algorithms*. Rochester, NY: Social Science Research Network, 2015. Disponível em: https://papers.ssrn.com/abstract=2660674. Acesso em: 27 maio. 2020.
[433] BUCHANAN, B. G. A. (Very) Brief History of Artificial Intelligence. *AI Magazine*, [*S. l.*], v. 26, n. 4, 2005. Disponível em: https://doi.org/10.1609/aimag.v26i4.1848. Acesso em: 21 jun. 2021.
[434] BURRELL, op. cit.

esse objeto.[435]

Os modelos mais comuns nessas aplicações são as redes neurais. Esses modelos são compostos por camadas. Uma camada de entrada, contendo, por exemplo, as imagens, a camada de saída, contendo a classificação dessas imagens, e as camadas intermediárias, que são capazes de calcular o peso de interconexões em elementos decompostos das imagens.

No entanto, as camadas intermediárias, por meio das quais o algoritmo reconhece uma imagem como pertencente ou não a uma categoria, não são expressas em termos lógicos, em categorias comparáveis aos que seres humanos utilizam para realizar a mesma classificação. Quando uma imagem é inserida no modelo, esse decompõe aquela em pequenos fragmentos compostos por poucos pixels que não correspondem a nenhuma forma conhecida ou comumente utilizada por seres humanos para explicar esse objeto.

Se as ações humanas são refletidas em uma práxis e compreendidas de uma forma simbólica, o processamento computacional funciona apenas com uma lógica algébrica. Tais modelos precisam ser interpretados para serem compreendidos, conforme discutido no trabalho de Burrell nos modelos que classificam *spam* nas caixas de *e-mail*. Um dos tipos de sistemas usados para filtro de *spam* são os *support vector machines* (SVMs). Consistem em uma forma de regressão usada para realizar classificações em variáveis binárias ou categóricas como sim ou não. No caso, *spam* ou não *spam*. A metodologia consiste em atribuir peso às palavras mais ou menos associadas com *spams*.

Embora a metodologia de atribuir pesos possa ser compreendida pelos humanos, é difícil de explicar como ou por que uma decisão sobre um *e-mail* em específico teve determinado resultado. Se podemos compreender que os algoritmos realizam classificações sobre a probabilidade de um e-mail ser um *spam* e possamos compreender quais palavras são mais ou menos relacionadas a esse incômodo, é difícil explicar por que um e-mail específico foi classificado como spam, visto que o peso de uma palavram em um contexto diferente, poderia ser outro.

Nesse sentido, é comum encontrarmos a terminologia relacionada a interpretabilidade dos modelos de *machine learning*. Dessa forma, para explicar os

---

[435] GOOGLE INC. ML Practicum: Image Classification. Google Developers. Disponível em: https://developers.google.com/machine-learning/practica/image-classification. Acesso em: 22 maio 2021.

resultados desses algoritmos é preciso desenvolver um conjunto de metodologias capaz de fornecer informações sobre o processamento e a relação entre algumas variáveis de acordo com alguns critérios de performance. No entanto, trata-se de uma aproximação do que o modelo realmente faz. Esse elemento remete novamente a discussão para a necessidade de desenvolverem-se mecanismos de explicação e governança.

4.2.1.3 Explicando modelos complexos de Inteligência Artificial

Os modelos científicos podem se configurar em diferentes tipos de explicações. Tentam explicar o que é um fenômeno, porque esse fenômeno ocorre, ou fornecem um quadro para sua compreensão sem que conheçamos todas as suas cadeias causais, em ramos da ciência onde ainda não há consenso ou teoria que permita a compreensão completa[436].

No campo científico, os modelos cumprem um papel em si mesmos. Por isso, quando utilizados em outros contextos, para serem compreendidos pelo público em geral, esses modelos carecem de um trabalho de tradução para outro tipo de explicação, mais cotidiana, que possa fornecer uma compreensão no nível necessário para que sua utilização possa ser instrumentalizada por não especialistas[437].

Avaliar a lógica dos algoritmos pode fornecer elementos para avaliarmos a justiça e o potencial discriminatório deles. No entanto, para explicar o “por que” de uma decisão, muitas vezes é preciso uma interpretação humana do algoritmo ou da sua lógica. Ou seja, a transposição da linguagem lógica de operação do computador para uma linguagem mais acessível. Como discutimos, a questão torna-se mais complexa quando se trata de explicar algumas classificações e sua dimensionalidade e as operações matemáticas envolvidas. E o problema se agrava quando se trata de explicar relações com várias dimensões, como é comum em contextos de *big data,* inteligência artificial e aprendizado de máquina.

As soluções para interpretabilidade desses sistemas envolvem simplificar o

---

[436] STREVENS, M. No understanding without explanation. *Studies in History and Philosophy of Science Part A*. v. 44, n. 3, p. 510–515, 2013. Disponível em: https://doi.org/10.1016/j.shpsa.2012.12.005. Acesso em: 21 jun. 2021.

[437] MITTELSTADT, B.; RUSSELL, C.; WACHTER, S. Explaining Explanations in AI. *In*: THE CONFERENCE, 2019, Atlanta, GA, USA. *Proceedings of the Conference on Fairness, Accountability, and Transparency — FAT'19*. Atlanta, GA, USA: ACM Press, 2019. p. 279–288. Disponível em: https://doi.org/10.1145/3287560.3287574. Acesso em: 27 maio 2020.

modelo ou desenvolver métodos para avaliar a capacidade de discriminação, correção dos dados ou dos seus resultados. Algumas abordagens tendem a criticar os algoritmos opacos de forma intencional e pedem por mais regulação ou transparência[438]. Outros defendem a auditoria dos algoritmos por agentes independentes. Outros argumentam que se pode educar setores amplos da sociedade capazes de avaliar o potencial discriminatório desses sistemas[439].

Referindo-se a algoritmos opacos em geral, Guidotti *et al.* apresentam uma metodologia para explicar uma decisão automática, a partir do estabelecido no GDPR como "meaningful explanation of the logic involved". A metodologia é chamada de LORE (*Local Rule-Based Explanation)*. Os pesquisadores realizaram testes no algoritmo do COMPAS e outras bases de dados para diferentes modelos e desenvolveram métricas para avaliar fidelidade e adequação dos modelos para explicar as decisões[440].

Os critérios estabelecem que essa explicação deve ser i) útil para que o sujeito possa compreender aquela decisão específica, de forma que devem ser locais e não globais; ii) ser apresentadas em uma linguagem mais próxima da lógica o quanto seja possível; e iii) que o algoritmo pode ser questionado quantas vezes for necessário para examinar sua lógica.

A metodologia fornece uma explicação agnóstica, que não faz menção ao algoritmo em questão, e se refere à análise dos *inputs* e *outputs* do sistema. Alguns modelos apresentam a lógica envolvida, na forma de explicações gerais que, no entanto, não explicam as decisões específicas. Os autores defendem que as explicações gerais não são significativas para os usuários afetados pelas decisões e por isso o mais apropriado seriam os modelos de explicação local como LIME[441].

Para a inteligência artificial, Mittelstadt *et al.,* em um *paper* recentemente publicado, apresentam uma revisão da literatura sobre explicabilidade e interpretabilidade, buscando convergências nas interpretações e apresentando os

---

[438] RUDIN, C. Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead. *ArXiv [cs, stat]*, [*S. l.*], 2019. Disponível em: http://arxiv.org/abs/1811.10154. Acesso em: 10 ago. 2020.

[439] DIAKOPOULOS, N. *Algorithmic accountability reporting:* on the investigation of black boxes. [*S. l.*]: Tow Center for Digital Journalism, 2014. Disponível em: https://academiccommons.columbia.edu/doi/10.7916/D8ZK5TW2. Acesso em: 21 jun. 2021.

[440] GUIDOTTI, R. *et al.* Local Rule-Based Explanations of Black Box Decision Systems. *ArXiv*, 2018. Disponível em: http://arxiv.org/abs/1805.10820. Acesso em: 30 jun. 2020.

[441] RIBEIRO, M. T.; SINGH, S.; GUESTRIN, C. "Why Should I Trust You?": Explaining the Predictions of Any Classifier. *ArXiv*, 2016. Disponível em: http://arxiv.org/abs/1602.04938. Acesso em: 25 ago. 2020.

principais problemas e lacunas[442]. Os autores afirmam que apesar do calor do debate, permanece confuso o significado do termo explicação. Em parte, porque as novas propostas no campo do *machine learning* contrastam com as definições estabelecidas em outras ciências humanas e no direito[443]. As explicações propostas por autores da área de *Machine Learning* se aproximam mais de modelos científicos do que de explicações cotidianas utilizadas pelas pessoas em contextos reais.

Além disso, tecem críticas aos modelos de explicação exclusivamente locais, visto serem enganadoras em nível global. Para o desenvolvimento de uma inteligência artificial confiável é preciso investir em interatividade e interpretabilidade que tornem mais fácil contestar decisões feitas por IA, inclusive possibilitando discussões sobre a justificação desses algoritmos e não apenas sobre uma decisão específica. Dessa forma, explicações locais deveriam ser incluídas num “explanation kit” mais amplo.

Os autores defendem o uso de explicações contrastivas das decisões, visto aproximarem-se das explicações cotidianas. Além disso, as explicações devem ser seletivas e socialmente significativas, relacionadas ao contexto. No entanto, para fugir dos problemas de explicações pontuais, é preciso pensar na relevância da resposta. É preciso que tal explicação contrafactual fornecida seja próxima o suficiente do caso específico e que o cenário hipotético descrito possa ocorrer no mundo real.

Watson e Floridi[444] apresentam uma discussão sobre a explicabilidade dos modelos de *machine learning*. No entanto, defendem que é necessário que o algoritmo seja interpretado em seu conjunto, para além da mera predição. Nesse sentido argumentam que as explicações podem variar em sua profundidade. Podem explicar o fenômeno (o resultado de uma predição algorítmica), as causas do fenômeno (fenômenos que têm mais ou menos influência na decisão) ou o próprio modelo a que tal fenômeno está relacionado (o algoritmo).

Os autores propõem um modelo de explicação tridimensional onde se considere a precisão, simplicidade e relevância da explicação. Esses critérios são estabelecidos através de cálculos pragmáticos dos agentes. Dessa forma, tornam-se pessoais, visto que dependem dos objetivos de cada agente. Concluem que a maior

---

[442] MITTELSTADT, B. D. *et al.* The ethics of algorithms: Mapping the debate. *Big Data & Society*, [*S. l.*], v. 3, n. 2, 2016. Disponível em: https://doi.org/10.1177/2053951716679679. Acesso em: 20 maio 2020.
[443] MILLER, T. Explanation in Artificial Intelligence: Insights from the Social Sciences. *ArXiv [cs]*, [*S. l.*], 2018. Disponível em: http://arxiv.org/abs/1706.07269. Acesso em: 14 ago. 2020.
[444] WATSON, D. S.; FLORIDI, L. The explanation game: a formal framework for interpretable machine learning. *Synthese*, 2020. Disponível em: https://doi.org/10.1007/s11229-020-02629-9. Acesso em: 8 jun. 2020.

disponibilidade de dados vai permitir um maior uso de inteligência artificial, e argumentam que eles só se tornarão mais transparentes caso pesquisadores se dediquem ao problema da explicação.

Os autores argumentam que as explicações devem ser a) precisas, capazes de refazer os procedimentos dos algoritmos, b) simples, de forma a serem compreendidas pelo agente que solicitou a explicação e c) relevantes, capazes de oferecer ao agente elementos que aumentem a sua capacidade de reagir à decisão. Nesse sentido, a explicação deve conter elementos globais e locais para atingir esses objetivos em cada contexto.

Há autores que defendem uma abordagem mais firme na defesa da transparência e do controle social e pedem maiores restrições ao uso de modelos opacos de *machine learning*, para que não sejam utilizados, por exemplo, em aplicações de alto risco[445]. Em muitos casos seria possível desenvolver um modelo interpretável e transparente com o mesmo grau de precisão, de forma que nas aplicações de alto risco deve-se exigir o uso de sistemas interpretáveis em vez de investir em modelos opacos que não fornecem explicações.

Black boxes podem derivar da incapacidade de compreensão humana (*deep learning*) ou de políticas corporativas. As explicações desses modelos aparecem então como replicações ou representações independentes descoladas do comportamento do sistema em questão. Consistem em discursos sobre como o sistema funciona, ou, na verdade, como deveria funcionar, e não como o modelo efetivamente funciona e sua relação com a realidade material. A discussão sobre explicabilidade, interpretabilidade, transparência podem estar perdendo o foco, que é a relação entre o sistema e as decisões, para uma discussão sobre a relação entre as explicações e os sistemas.

Rudin aponta alguns problemas recorrentes que aparecem nas discussões sobre o tema. Argumenta que há um falso *trade off* entre interpretabilidade e precisão. Essa relação não se observa na prática e o sistema pode ser aprimorado conforme é reprogramado, sendo possível torná-lo mais simples. Mesmo nos modelos mais complexos de visão computacional, onde *deep learning* traz ganhos de eficiência, é possível inserir parâmetros de interpretabilidade no sistema sem perder precisão. Os

[445] RUDIN, C. Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead. *ArXiv [cs, stat]*, [*S. l.*], 2019. Disponível em: http://arxiv.org/abs/1811.10154. Acesso em: 10 ago. 2020.

algoritmos impossíveis de interpretação são muito úteis em contextos de descobertas, mas não são o objetivo final naquelas aplicações. Na realidade o objetivo é interpretar tais resultados como forma de contribuir em determinada ciência.

Além disso, argumenta que as explicações fornecem esclarecimentos que não são fiéis ao processo que efetivamente ocorreu dentro de um modelo. As explicações, por serem simplificações, podem não ser reais para alguns casos ou resultados do modelo. A interpretação correta depende de expertise e capacidade de conhecer profundamente as limitações e possibilidades de erro dos modelos utilizados em uma decisão, e nada garante que as explicações forneçam tais elementos.

Sobre a relevância, a autora alerta que explicações nem sempre são claras ou fornecem informações suficientes sobre o que o sistema realmente faz. Elas podem deixar tantas informações de fora que o modelo deixa de fazer sentido. As explicações, além disso, focam no funcionamento correto do algoritmo ou ideal do algoritmo, não no seu funcionamento real.

Ela utiliza o exemplo dos modelos de cálculo de risco no sistema penal. Os algoritmos *black boxes* não são compatíveis com modelos nos quais informações fora da base de dados precisam ser inseridas na análise de risco. As condições reais do crime devem impactar além das categorias previstas no sistema, no entanto, não é possível ao juiz calibrar com base nessas condições exteriores a base de dados.

Outro argumento muito relevante diz respeito a como o uso desses algoritmos pode levar a maiores chances de erros humanos. O sistema pode estar errado ou com defeito, e não é possível avaliar isso com base nas explicações. Se em alguns casos podem ser preferíveis sistemas opacos com explicações, esses casos são minoria. A autora alerta que o direito à explicação não vai garantir que essa explicação seja suficiente ou mesmo correta. Antes de aceitar o uso de modelos opacos as autoridades devem insistir e incentivar o desenvolvimento de modelos interpretáveis para aplicações de alto risco, e restringir o uso de *black boxes* a poucos casos com baixo risco.

Rudin ainda rebate o argumento de que os modelos *black box* são importantes para evitar a prática de manipulação *(gaming*). Em primeiro lugar porque a possibilidade de ser manipulado pode indicar falhas em seu desenho. A prática de manipulação só é possível quando os indicadores utilizados para *rankeamento* não correspondem a elementos relevantes em si mesmo para o que se está avaliando. Argumenta que sistemas de *rankeamento* que se utilizem de critérios objetivos de

qualidade podem ter um impacto positivo ao incentivar o desenvolvimento dos parâmetros verdadeiros de qualidade e não meros números. E em alguns casos o argumento não faz sentido visto que a possibilidade de mudar o comportamento para manipular o *output* do sistema é bom em si mesmo como nos *scores* de crédito.

O outro argumento enfrentado por Rudin é de que as explicações contrafactuais seriam suficientes. Essa abordagem é limitada visto que não descreve como funciona o sistema. Essa informação mínima para apresentar novos resultados (se você tivesse uma dívida $1000 menor, você estaria autorizado para o empréstimo) é diferente para cada pessoa, visto que combina informações fora da base de dados.

O *paper* então apresenta os maiores entraves para o desenvolvimento desses sistemas. O principal deles são as necessidades corporativas. Os modelos transparentes não oferecem possibilidades de lucro. Além disso, podem custar caro, visto que a interpretabilidade coloca alguns elementos adicionais no sistema, que tornam mais difícil de resolver um problema. No entanto, para modelos de aplicações de alto risco, a autora argumenta que esse custo se torna justificável, visto que sistemas não interpretáveis podem trazer problemas e custos extras para aplicação.

A autora reconhece que o GDPR estabelece um direito à explicação. No entanto, não requer um modelo interpretável. A legislação estabelece apenas a obrigação de explicação, sem, contudo, estabelecer que essa explicação seja precisa, completa, verdadeira, confiável em relação ao modelo utilizado. Explicações insatisfatórias podem contornar essa disposição. A autora propõe mecanismos mais rígidos como estipular que as companhias não possam utilizar modelos opacos quando um modelo interpretável estiver disponível.

Os méritos do trabalho de Rudin são indiscutíveis em apresentar os limites da explicação de sistemas opacos e da necessidade de se discutir a realidade do sistema para atingirmos maior *accountability*. No entanto, o trabalho incide no erro apontado por Miller, de trabalhos que ignoram as perspectivas das ciências sociais sobre os processos reais de compartilhamento de informações entre indivíduos em determinada sociedade[446]. Mesmo os modelos interpretáveis ainda apresentam uma limitação visto que ainda assim não serão automaticamente compreendidos pelas pessoas que possam por eles serem afetadas. Assim como algoritmos mais simples e não opacos, eles ainda dependem de uma interpretação e, para serem

[446] MILLER, T. Explanation in Artificial Intelligence: Insights from the Social Sciences. *ArXiv [cs]*, [*S. l.*], 2018. Disponível em: http://arxiv.org/abs/1706.07269. Acesso em: 14 ago. 2020.

compreendidos pelas pessoas comuns, de uma explicação contextual.

Para garantir o efetivo direito previsto nas legislações de proteção de dados sobre as decisões automatizadas, é importante levar em consideração o nível de conhecimento do receptor da explicação, suas necessidades e a sua capacidade de autodeterminação. Nesse sentido, a correção dos modelos e sua precisão não são suficientes. Dessa forma a explicação torna-se também uma forma de justificação, como apontado no trabalho de Mittelstad *et al*.[447]. Para uma ação ser justa e legítima, é preciso que um indivíduo seja capaz de compreender as decisões feitas sobre ele e ser capaz de questioná-las.

4.2.2 Limitações associadas à capacidade de compreensão do titular de dados pessoais

Além das estratégias corporativas e dos limites técnicos tratados anteriormente, outra forma de opacidade de algoritmos também mencionada por Jenna Burrell[448] diz respeito à capacidade de compreensão da população não especialista. Os algoritmos não são conhecidos por todos e dependem de certa compreensão sobre linguagens computacionais e matemáticas. Dessa forma, um dos limites ao direito à explicação consiste nessa barreira entre o funcionamento de uma tecnologia, ainda que transparente, e capacidade de compreensão do público amplo, usuário ou afetado por ela.

Já discutimos como a noção de explicação implica em uma relação bilateral que depende da assimilação do receptor. A explicação, dessa forma, só é efetiva quando tem como resultado a compreensão da outra parte[449]. A questão que se coloca, portanto, diz respeito a formas de construir sistemas capazes de se fazerem compreensíveis para todos os tipos de público afetados por decisões automatizadas.

Motores de automóveis, turbinas de avião, processos químicos de tratamento de água, exames de DNA e previsões do tempo também são produtos de sistemas

---

[447] MITTELSTADT, B.; RUSSELL, C.; WACHTER, S. Explaining Explanations in AI. *In*: THE CONFERENCE, 2019, Atlanta, GA, USA. *Proceedings of the Conference on Fairness, Accountability, and Transparency — FAT'19*. Atlanta, GA, USA: ACM Press, 2019. p. 279–288. Disponível em: https://doi.org/10.1145/3287560.3287574. Acesso em: 27 maio 2020.

[448] BURRELL, J. *How the Machine "Thinks:" Understanding Opacity in Machine Learning Algorithms*. Rochester, NY: Social Science Research Network, 2015. Disponível em: https://papers.ssrn.com/abstract=2660674. Acesso em: 27 maio. 2020.

[449] STREVENS, M. No understanding without explanation. *Studies in History and Philosophy of Science Part A*. v. 44, n. 3, p. 510–515, 2013. Disponível em: https://doi.org/10.1016/j.shpsa.2012.12.005. Acesso em: 21 jun. 2021.

sociotécnicos. Para a maior parte da população há também um desconhecimento dessas tecnologias, pois são matérias para especialistas. A questão, contudo, tem um aspecto mais grave em relação às tecnologias de informação pois estas podem afetar diretamente nossos direitos e a nossa personalidade[450].

As tecnologias da informação estão presentes no cotidiano de grande parte da população. No entanto, a universalização do acesso não significa, em si, a capacidade de compreender o que é uma tecnologia. Com o desenvolvimento de interfaces para usuários não especialistas, é possível utilizar produtos e serviços sem compreender o seu funcionamento, seus riscos, seus processos de criação e seu papel dentro de sistemas econômicos, políticos e sociais.

O desconhecimento de uma área específica da ciência aplicada não torna, necessariamente, uma pessoa excluída, embora o acesso à formação técnica avançada esteja constantemente associada a melhores rendimentos e status sociais. Contudo, no contexto de ubiquidade das tecnologias da informação, a falta de familiaridade com computadores e códigos são fatores de vulnerabilidade que podem agravar desigualdades sociais e afetar até mesmo alguns pilares da sociedade democrática.

Em resposta a essa questão, os governos e a sociedade civil organizada tem se preocupado com a difusão do conhecimento sobre a sociedade digital. O ensino de tecnologias de computação é muitas vezes incorporado aos currículos educacionais como forma de fomentar o desenvolvimento econômico a partir de habilidades tecnológicas, sociais e cognitivas para um novo modo de organização da produção e da sociedade[451].

Uma primeira abordagem consiste em pensar na “alfabetização” tecnológica em termos de acesso a computadores e à internet. Embora o acesso à Internet esteja sendo cada vez mais ampliado e difundido, ainda há um longo caminho a ser percorrido para a garantia do acesso universal. Os problemas enfrentados no Brasil para distribuição do auxílio emergencial fornecido pelo governo no contexto da pandemia da Covid-19 demonstram como o acesso à internet ainda não está

---

[450] CHENEY-LIPPOLD, J. *We are data*: algorithms and the making of our digital selves. New York: New York University Press, 2017.

[451] DAVIES, R. S. Understanding Technology Literacy: A Framework for Evaluating Educational Technology Integration. *TechTrends*, [*S. l.*], v. 55, n. 5, 2011. Disponível em: https://doi.org/10.1007/s11528-011-0527-3. Acesso em: 21 jun. 2021. p. 45.

disponível a todo mundo e como essa escassez pode impactar direitos básicos.[452]

Superar a questão do acesso, contudo, não implica que a tecnologia seja compreendida e utilizada em todo o seu potencial. Autores argumentam que é importante levar em consideração o contexto em que as pessoas usam a tecnologia. Embora o acesso a *smartphones* de fato tenha aumentado nos últimos anos até mesmo nas menores faixas de renda, sua utilização é limitada a poucos aplicativos. Dessa forma, embora com acesso à internet e com *smartphones*, muitas pessoas não possuíam familiaridade e facilidade para realizar configurações ou contornar *bugs* do aplicativo.

Essa desigualdade no acesso e uso de tecnologias digitais é um problema que diversos autores trabalharam a partir de uma questão apelidada de *digital divide*[453]. Não é simples explicar a divisão, suas causas, consequências e formas de superação. A começar pela variação do acesso e conhecimento em função da idade. É comum percebermos os mais velhos como menos familiarizados com tecnologias digitais[454]. Mas essa menor habilidade pode ser explicada tanto por características físicas, decorrentes das características cognitivas dos mais velhos, ou geracionais, decorrentes da experiência com o seu uso de tecnologias ser mais difundida entre os mais novos.

Outro problema enfrentado diz respeito à variação em função da renda. O aumento da renda pode significar mais familiaridade com tecnologias. Em um primeiro momento podemos afirmar que isso se deve ao acesso mais facilitado às tecnologias. Contudo, visto as transformações no mercado de trabalho, a falta de conhecimentos sobre tecnologias digitais pode constituir também uma causa da desigualdade de renda, e não sua consequência[455].

---

[452] FAMÍLIAS sem acesso à internet não conseguem usar o dinheiro do auxílio emergencial. *Jornal Nacional*, 9 abr. 2021. Disponível em: https://g1.globo.com/jornal-nacional/noticia/2021/04/09/familias-sem-acesso-a-internet-nao-conseguem-usar-o-dinheiro-do-auxilio-emergencial.ghtml. Acesso em: 22 maio 2021.

[453] CHEN, W.; BARRY, W. Charting Digital Divides: Comparing Socioeconomic, Gender, Life Stage, and Rural-Urban Internet Access and Use in Five Countries. *In*: DUTTON, W. H.; KAHIN, B.; O'CALLAGHAN, R.; WYCKOFF, A. W. *Transforming Enterprise:* The Economic and Social Implications of Information Technology. MITP, 2004. p. 467–497. *E-book*.

[454] LOSH, S. C. Generation versus aging, and education, occupation, gender and ethnicity effects in U.S. digital divides. *In: 2009 Atlanta Conference on Science and Innovation Policy,* Atlanta, 2009. p. 1–8. Disponível em: https://doi.org/10.1109/ACSIP.2009.5367820. Acesso em: 21 jun. 2021.

[455] TOMCZYK, L. *et al.* Digital Divide in Latin America and Europe: Main Characteristics in Selected Countries. *In*: IBERIAN CONFERENCE ON INFORMATION SYSTEMS AND TECHNOLOGIES (CISTI), 14., 2019. Disponível em: https://doi.org/10.23919/CISTI.2019.8760821. Acesso em: 21 jun. 2021.

Podemos apontar, ainda, causas históricas e geracionais, visto que os primeiros usuários das tecnologias digitais eram homens, ricos, bem-educados, brancos e de países desenvolvidos[456]. Os países do sul foram retardatários na adoção dessas tecnologias, de forma que é de se esperar que isso se reflita nas habilidades da população entre os países e dentro dos próprios países, visto que há diferenças entre classes sociais, raça e gênero.

Não seria exagerado supor que essa falta de diversidade entre os primeiros usuários esteja relacionada à implantação de aplicações com vieses racistas, sexistas ou xenófobos. A divisão desigual[457] entre o poder computacional é apontada como um dos fatores para consolidação desse tipo de viés. Nos últimos anos temos observado estratégias corporativas de implementação de políticas de diversidade como forma de incrementar os resultados de suas equipes[458]. No entanto, esse processo está longe de chegar a um fim, de forma que devemos nos confrontar com o problema da falta de diversidade por mais alguns anos.

Se a diversidade cumpre um papel central na produção de tecnologias, principalmente de algoritmos, é importante levar em consideração, além disso, como a diferença no nível de conhecimento sobre algoritmos pode impactar de forma diferentes algumas minorias, potencializando desigualdades estruturais, enquanto alvos ou objetos de processamentos de dados.

Mais do que colocar as mãos no teclado e ter acesso à internet, a proficiência nas tecnologias envolve utilizá-la de forma significativa, consciente de seus objetivos, riscos e limites. É preciso compreender como operam os fluxos informacionais e como os dados pessoais podem produzir efeitos em nossas realidades para que os agentes possam, de maneira racional, adequar o seu comportamento.

Os algoritmos, principalmente no campo da inteligência artificial, são parametrizações e quantificações de aspectos sociais ou naturais que carregam em si premissas e uma visão de mundo. Por isso requerem certo nível de conhecimento e expertise no seu uso e interpretação de seus resultados para além da mera noção

---

[456] CHEN, W.; BARRY, W. Charting Digital Divides: Comparing Socioeconomic, Gender, Life Stage, and Rural-Urban Internet Access and Use in Five Countries. *In*: DUTTON, W. H.; KAHIN, B.; O'CALLAGHAN, R.; WYCKOFF, A. W. *Transforming Enterprise:* The Economic and Social Implications of Information Technology. MITP, 2004. p. 467–497. *E-book*.

[457] AHMED, N.; WAHED, M. The De-democratization of AI: Deep Learning and the Compute Divide in Artificial Intelligence Research. *ArXiv*, 2020. Disponível em: http://arxiv.org/abs/2010.15581. Acesso em: 9 nov. 2020.

[458] ROCK, D.; GRANT, H. Why Diverse Teams Are Smarter. *Harvard Business Review*, 4 nov. 2016. Disponível em: https://hbr.org/2016/11/why-diverse-teams-are-smarter. Acesso em: 27 out. 2020.

propagada de que se tratam de resultados objetivos. Uma noção sobre os dados e sobre metodologias científicas é essencial para a compreensão dos resultados.[459]

Essa falta de conhecimento pode afetar significativamente pessoas e a sociedade como um todo. Os mecanismos de busca oferecem um vasto conteúdo disponível na rede. Contudo, ao não conhecer como funciona o ranqueamento e a indexação, pessoas podem interpretar seus resultados de forma equivocada. Os algoritmos podem fazer um excelente trabalho de curadoria de conteúdos nas plataformas de mídia. No entanto, ao desconhecerem como funcionam, as pessoas podem estar mais suscetíveis à propagação de notícias falsas, discursos de ódio ou propagandas enganosas. Nas análises de crédito, análises para planos de saúde ou testes profissionais, o desconhecimento sobre a forma de operação dos algoritmos pode fazer com que um sujeito não seja capaz de identificar erros ou vieses que o prejudiquem[460].

Pela ubiquidade desses sistemas e sua opacidade no contexto atual, o conhecimento sobre algoritmos já é apontado como uma competência necessária para os currículos escolares como forma de promover a cidadania e maior controle sobre a própria personalidade num futuro que será ainda mais permeado pela inteligência artificial[461].

Diante da desigualdade das competências e conhecimentos sobre tecnologia, a capacidade de explicação torna-se limitada pelo nível de consciência do público a que se destina. E é nesse sentido que o reconhecimento do direito à explicação torna-se essencial. Se os sujeitos não conhecem sobre a tecnologia, o que ela faz, por que o faz e como funciona, não serão, consequentemente, capazes de compreender os riscos e de se preocupar com medidas de segurança, de mitigação ou de questionamento. O mero fornecimento de informações não estabelece o dever de considerar a comunicação como uma relação bidimensional onde a compreensão seja um requisito para sua validade[462].

---

[459] KOLKMAN, D. The (in)credibility of algorithmic models to non-experts. *Information, Communication & Society*, v. 0, n. 0, p. 1–17, 2020. Disponível em: https://doi.org/10.1080/1369118X.2020.1761860. Acesso em: 28 jun. 2021.

[460] BAKER, J. J. Beyond the Information Age: The Duty of Technology Competence in the Algorithmic Society. South Carolina Law Review, v. 69, n. 557, 2018. Disponível em: https://ttu-ir.tdl.org/handle/2346/73913. Acesso em: 21 out. 2020.

[461] PEW RESEARCH CENTER: INTERNET, SCIENCE & TECH. *Code-Dependent*: Pros and Cons of the Algorithm Age. [*S. l.*]: Pew Research Center, 2017. p. 74 et seq.

[462] WATSON, D. S.; FLORIDI, L. The explanation game: a formal framework for interpretable machine learning. *Synthese*, 2020. Disponível em: https://doi.org/10.1007/s11229-020-02629-9. Acesso em: 8 jun. 2020.

O campo de estudo conhecido como *human computer interactions* (HCI) é essencial para entender essas questões. O conceito ganha relevo com o surgimento dos primeiros computadores e diz respeito à forma como dois sistemas podem interagir: de um lado, os sistemas computacionais e, de outro, o sistema representado pelo pensamento humano.[463] O desenvolvimento de interfaces de usuários e de aplicações na internet representou um marco importante no campo desses estudos. Além disso, a multiplicação de usuários de computadores pessoais e *smartphones* trouxe novos desafios[464].

O desafio de construir interfaces consiste em formas de estabelecer comunicações e interações eficientes. O *design* do sistema depende do conhecimento sobre o sistema e o usuário. No contexto de decisões automatizadas, consiste em permitir a compreensão do sistema, quanto é utilizado, e possibilitar que os usuários se informem sobre o funcionamento e sobre os seus resultados.

Os usuários podem ser especialistas em determinada área ou podem ser pessoas comuns submetidas a algum processo decisório realizado por algoritmos, o que envolve diferentes estratégias de comunicação. Os desenvolvedores de aplicações defrontam-se com o desafio de construir um sistema que seja compreensível e utilizável por milhões de usuários. Por óbvio, não é possível desenhar um modelo de interação que leve em consideração cada usuário individual, visto que cada pessoa possui um nível de conhecimento diferente sobre o tema.

Na concepção de modelos de explicação, sejam eles automatizados ou fornecidos por seres humanos, é preciso analisar as necessidades do usuário, o propósito da tarefa e a melhor forma de estruturar a interação[465]. É preciso que tal modelo seja centrado nos seres humanos em sua operacionalidade, com aparência coerente e fácil de compreender, bem como capaz de fornecer opções de busca e um mecanismo de ajuda realmente efetivo.

Pensar nas necessidades dos usuários e compreender seu nível de competências sobre determinados temas não é tarefa fácil. Há uma série de

---

[463] SINHA, G.; SHAHI, R.; SHANKAR, M. Human Computer Interaction. *In*: INTERNATIONAL CONFERENCE ON EMERGING TRENDS IN ENGINEERING AND TECHNOLOGY, 3., 2010. p. 1–4. Disponível em: https://doi.org/10.1109/ICETET.2010.85. Acesso em: 21 jun. 2021.

[464] KAIYAN, N. Exploratory study of implicit theories in human computer interaction. *Proceedings Sixth Australian Conference on Computer-Human Interaction*. [*S. l.: s. n.*], 1996. p. 338–339. Disponível em: https://doi.org/10.1109/OZCHI.1996.560158. Acesso em: 21 jun. 2021.

[465] CHAO, G. Human-Computer Interaction: Process and Principles of Human-Computer Interface Design. *In*: *2009 International Conference on Computer and Automation Engineering*, [*S. l.*]: 2009. p. 230–233. Disponível em: https://doi.org/10.1109/ICCAE.2009.23. Acesso em: 21 jun. 2021.

metodologias para desenvolver uma noção de um usuário que seja representativo do público a que se destina a explicação. É possível realizar trabalhos etnográficos ou estudos de cenário, elencando as possíveis necessidades[466]. É importante destacar que há peculiaridades e diferentes públicos. Há usuários novatos e *experts*. Há usuários familiarizados com interfaces, com algoritmos, com estatística ou determinado campo da ciência, enquanto há outros para os quais esses temas não são tão claros.

Além disso, mesmo entre *experts* e novatos, há intenções diferentes para o uso do sistema, que, se não forem compreendidas, podem tornar uma explicação não efetiva. Quando há uma decisão, dificilmente uma pessoa está curiosa sobre todos os seus aspectos. A aplicação precisa permitir que a pessoa compreenda de quais informações necessita.

Nesse sentido, o direito à explicação demanda que as organizações desenvolvam explicações, e não apenas uma explicação. Um texto, vídeo ou material que seja coerente e muito exato quanto ao processo ao qual se refere não será uma explicação a menos que seja efetivamente compreendido. Por isso, mecanismos de *feedback* são essenciais para identificar se os sujeitos foram capazes de compreender a intencionalidade, os usos, os objetivos da tecnologia, bem como os critérios de determinadas decisões. Então, a depender das questões envolvidas, é preciso explicar as bases de dados utilizadas para o desenvolvimento dos sistemas, as métricas de eficiência, as metodologias de avaliação e monitoramento etc. Caso contrário, haverá um verdadeiro esvaziamento do seu direito à explicação, que não passará de uma ficção jurídica, e efetivamente lhe impedirá o exercício de outros direitos correlatos, como o de revisão de tais decisões.

4.3 DESAFIOS INSTITUCIONAIS

A efetiva garantia e implementação do direito à explicação depende em larga medida das capacidades e dos arranjos institucionais voltados ao *enforcement* da legislação de proteção de dados pessoais. Nesse sentido, conforme assinalam Raab e Szekely, a eficácia da regulação de proteção dados pessoais depende muito mais

[466] FISCHER, G. User Modeling in Human–Computer Interaction. *User Modeling and User-Adapted Interaction*, v. 11, n. 1, p. 65–86, 2001. Disponível em: https://doi.org/10.1023/A:1011145532042. Acesso em: 21 jun. 2021.

da forma como as autoridades de *enforcement* desempenham suas tarefas e aplicam o texto da lei do que das potencialidades ou déficits da legislação em si.[467] Se, de um lado, a existência de capacidades e arranjos institucionais é crucial para a garantia dos direitos dos titulares, a ausência de tais capacidades constitui relevante barreira à sua efetiva garantia e implementação. Nesta seção, buscaremos discutir como as autoridades de proteção de dados e instâncias judiciais têm atuado para garantir o *enforcement* da regulação, analisando como essa competência tem sido exercida na prática, a partir das limitações e condições concretas de atuação desses órgãos.

Na subseção 4.3.1, analisaremos os desafios institucionais enfrentados pelas autoridades nacionais de proteção de dados, pelas autoridades supervisoras e pelas autoridades subnacionais no cenário europeu. Essas entidades têm, dentre outras, a função de realizar o *enforcement* do Regulamento e das leis nacionais e subnacionais de proteção de dados, o que inclui a tarefa de garantir e verificar se os direitos dos titulares estão sendo garantidos, dentre os quais o direito à explicação. Todavia, em face da complexidade dos sistemas algorítmicos, de limitações financeiras e de recursos humanos enfrentadas pelas autoridades, bem como de outros fatores, nem mesmo estes entes dispõem de capacidade e dos instrumentos técnicos e regulatórios necessários para desempenhar satisfatoriamente suas funções. Neste contexto, buscaremos analisar como essas instituições têm se comportado diante de tais limitações, a partir de uma análise da performance e dos desafios enfrentados nas últimas três décadas, tomando como ponto de partida a promulgação da Diretiva 95/46/CE. De antemão, cabe-nos destacar que esta seção não possui a pretensão de esgotar a análise dos desafios institucionais enfrentados pelos agentes de *enforcement*, que são inúmeros e dos mais diversos, mas tão somente destacar alguns daqueles que impõem certas limitações à plena garantia do direito à explicação.

4.3.1 Entidades supervisoras

A partir da análise geracional das leis de proteção de dados pessoais apresentada por Mayer-Schöenberger, discutida no Capítulo 2 do presente trabalho,

[467] RAAB, C.; SZEKELY, I. Data Protection Authorities and Information Technology. *Computer Law and Security Review*, 1 jul. 2017. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2994898. Acesso em: 14 maio 2021.

é possível situar a disseminação de autoridades supervisoras independentes no contexto da 4ª geração de leis de proteção de dados pessoais. De acordo com o autor, a partir da década de 1990, há a emergência de uma quarta geração de leis de proteção de dados, marcada pela conjugação da tutela individual e de mecanismos de tutela coletiva, pela coexistência de leis gerais com leis setoriais, bem como pelo fortalecimento do aparato institucional de *enforcement*, com a disseminação de autoridades independentes.[468] Conforme assinala Wimmer, contudo, apesar de a primeira norma exigindo a criação desse tipo de autoridade só ter surgido em meados da década de 1990, com a Diretiva 95/46/CE, "[...] àquela altura um grande número de países europeus já contava com autoridades dessa natureza em funcionamento."[469] Neste sentido, a previsão desse tipo de órgão está presente desde as primeiras leis sobre a matéria, como a lei do estado alemão de Hesse, de 1970, que previa uma estrutura administrativa responsável pelo *enforcement*. Posteriormente, esse modelo de leis acompanhadas de autoridades de *enforcement* foi se fortalecendo na Europa após a instituição da autoridade francesa, a CNIL, e, posteriormente, com a Diretiva 95/46/CE, de 1995, que tornou obrigatória a instituição de uma ou mais autoridades de controle em cada Estado-Membro da UE. Pouco tempo depois, já no início dos anos 2000, esse modelo foi reforçado pela Carta de Direitos Fundamentais da UE, que prevê que a constituição de uma autoridade "[...] é ponto integral e orgânico do próprio direito fundamental à proteção de dados pessoais."[470] Ao suceder a Diretiva, a GDPR incorporou esse modelo de autoridades, tendo inovado ao prever mecanismos de cooperação entre essas autoridades no âmbito do bloco. A previsão dessas autoridades, contudo, não é característica do modelo europeu: cerca de 80% das 132 leis gerais de proteção de dados promulgadas até 2019 contavam com autoridades supervisoras em operação.[471]

---

[468] MAYER-SCHOENBERGER, V. Generational development of data protection in Europe. *In*: AGRE, P. E.; ROTENBERG, M. (Eds.). *Technology and Privacy:* The New Landscape. Cambridge: MIT Press, 1998. p. 219-241.

[469] WIMMER, M. Autoridades de proteção de dados pessoais no mundo: fundamentos e evolução na experiência comparada. *In:* PALHARES, Felipe (Coords.). *Temas atuais de proteção de dados*. São Paulo: Thomson Reuters, 2020. p. 157.

[470] DONEDA, D. A Autoridade Nacional de Proteção de Dados e o Conselho Nacional de Proteção de Dados. *In*: MENDES, L. S.; DONEDA, D.; SARLET, I. W.; RODRIGUES JÚNIOR, O. L.; BIONI, B. (Coords.). *Tratado de Proteção de Dados Pessoais*. Rio de Janeiro: Forense, 2021. p. 460; No mesmo sentido, cf.: WIMMER, M. Autoridades de proteção de dados pessoais no mundo: fundamentos e evolução na experiência comparada. *In:* PALHARES, Felipe (Coords.). *Temas atuais de proteção de dados*. São Paulo: Thomson Reuters, 2020. p. 157.

[471] Ibidem, p. 461.

Um dos principais desafios institucionais enfrentados pelas autoridades supervisoras com potencial de comprometer a garantia e a satisfatória implementação do direito à explicação no cenário europeu consiste na ausência de conhecimento técnico adequado sobre questões ligadas às novas Tecnologias da Informação e Comunicação. Considerando que o direito à explicação se dá no contexto de decisões automatizadas e operacionalizadas por meio de novas TICs, dentre as quais o *Big Data*, a Inteligência Artificial e o *machine learning*, a ausência de *expertise* técnica sobre essas novas tecnologias pode comprometer a atuação das autoridades supervisoras e limitar, direta ou indiretamente, a efetiva garantia e implementação do direito à explicação.

De acordo com Raab e Szekely, as autoridades supervisoras podem ser definidas como instituições *multi-task*, uma vez que acumulam uma série de funções e competências. Conforme assinalam os autores, essas funções e competências não são desempenhadas da mesma forma e com a mesma intensidade pelas diferentes autoridades: cada instituição pode priorizar determinadas funções (*advocacy* em detrimento de *enforcement* sancionatório, por exemplo), o que nos leva a um cenário de coexistência de múltiplas autoridades com diferentes perfis e abordagens de atuação. Os autores destacam, contudo, que o que é comum às diferentes autoridades é a sua necessidade de entender como as tecnologias da informação e comunicação trazem implicações para a proteção de dados pessoais e da privacidade. Conforme argumentam, é possível afirmar, com base na literatura, que as poucas autoridades existentes até a década de 1980 contavam com quadros técnicos dotados de conhecimento adequado para entender as implicações trazidas para a garantia da privacidade e a proteção de dados pessoais pelas novas tecnologias.[472] Todavia, os autores se questionam sobre se seria ainda este o cenário atualmente, no contexto de novos desenvolvimentos tecnológicos e dos crescentes desafios para sua compreensão impostos às diversas autoridades de supervisão europeias.[473]

---

[472] RAAB, C.; SZEKELY, I. Data Protection Authorities and Information Technology. *Computer Law and Security Review*, 1 jul. 2017. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2994898. Acesso em: 14 maio 2021. p. 3.

[473] "Perhaps that was true of the relatively few DPAs at that time in the larger countries of the EU, but is it still the case today, when the technological explosion, the proliferating demands placed upon DPAs, and the growth in their numbers across the EU at national and sub-national levels cast some doubt on their ability to deploy such knowledge in their activities 'on the ground'?" (RAAB, C.; SZEKELY, I. Data Protection Authorities and Information Technology. *Computer Law and Security Review*, 1 jul. 2017. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2994898. Acesso em: 14 maio 2021. p. 3).

Partindo desta hipótese, Raab e Szekely conduziram uma pesquisa empírica com 79 autoridades regionais, nacionais e subnacionais de proteção de dados com o objetivo de analisar em que medida as DPAs estavam a par das recentes mudanças e desenvolvimentos provocados pelas TICs com impacto para a proteção de dados pessoais e como essas mudanças impactavam suas atividades regulatórias:

> This survey aimed to find out the extent to which DPAs were abreast of changes in ICTs with which personal data are processed and which powerfully shape the terrain on which DPAs' regulatory and supervisory activities take place. *The authors' interest in this subject was fuelled by a perception that DPAs – among other shortcomings – were particularly deficient in their understanding of ICTs, so that their ability to regulate information processing would be compromised by deficiencies in their comprehension of technological changes that had important consequences for the protection of rights to privacy and data protection*.[474] (grifo nosso).

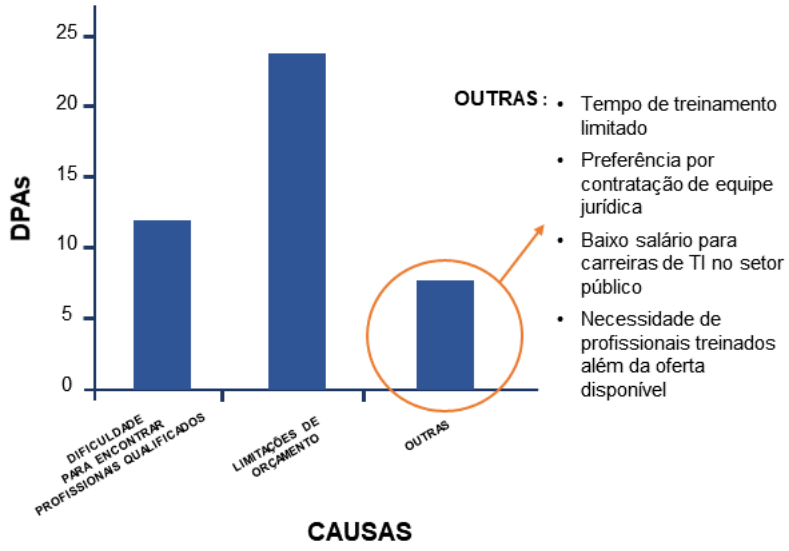
A pesquisa conduzida pelos autores levou a resultados bastante interessantes, que permitem compreender como a limitação de conhecimento técnico pode comprometer o *enforcement* das autoridades supervisoras em determinadas questões e, em última análise, a garantia e implementação de determinados direitos intimamente relacionados às novas tecnologias de processamento de dados, como é o caso do direito à explicação.

Um dos primeiros achados da pesquisa a serem destacados diz respeito à percepção dos respondentes, membros do corpo técnico das DPAs, sobre o nível de conhecimento técnico das autoridades sobre questões relacionadas às novas TICs. A pesquisa conduzida incluía uma pergunta sobre como as autoridades de proteção de dados avaliavam o nível de experiência e conhecimento das DPAs em geral na área de tecnologias de informação e comunicação. De acordo com os autores, a maioria dos funcionários entrevistados classificou a especialização das DPAs em geral em nível médio (entre 5 e 7 em uma escala numérica de 1 a 10), e nenhuma autoridade avaliou o nível de especialização das DPAs no nível mais baixo, tendo havido apenas um entrevistado que atribuiu a nota mais alta. Aqueles que atribuíram as notas mais baixas atribuíram o déficit de *expertise* a diferentes causas, das quais duas receberam maior destaque: a limitação de recursos financeiros, em primeiro lugar, seguida da

[474] RAAB, C.; SZEKELY, I. Data Protection Authorities and Information Technology. *Computer Law and Security Review*, 1 jul. 2017. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2994898. Acesso em: 14 maio 2021. p. 3-4.

dificuldade de encontrar pessoal qualificado, conforme ilustrado na figura abaixo:

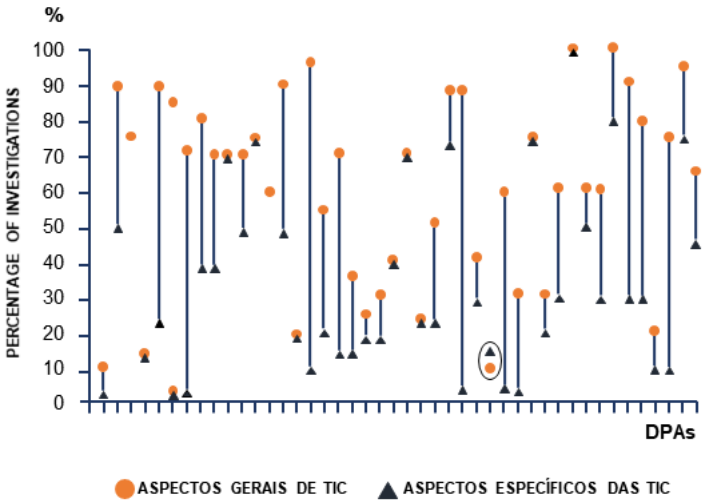
GRÁFICO 1 – CAUSAS DO DÉFICIT DE *EXPERTISE* INTERNA NAS DPAS EUROPEIAS



Fonte: RAAB; SXEKELY, 2017,p.8.

Uma das questões presentes no estudo buscava ainda identificar a frequência de investigações relacionadas a TICs em geral e a frequência de investigações que demandavam conhecimento e *expertise* específicos em TICs. As respostas foram sumarizadas no quadro abaixo. No eixo X estão representadas as DPAs respondentes e no eixo Y a porcentagem de investigações. Os círculos representam a porcentagem de investigações envolvendo aspectos gerais das TICs, enquanto os triângulos representam a porcentagem de investigações envolvendo *expertise* em aspectos específicos das TICs:

GRÁFICO 2 – Investigações das dpas consultadas envolvendo aspectos Gerais das tics e *expertise* específica em aspectos relacionados às tics



Fonte: RAAB; SXEKELY, 2017,p.9.

Dado o baixo grau de *expertise* das DPAs em questões técnicas envolvendo TICs e o alto número de investigações envolvendo questões técnicas, sejam elas relacionadas a aspectos gerais ou específicos, o estudo considerou ser moderado o risco de controladores, em especial aqueles na posição de provedores de serviço, induzirem as DPAs ao erro em matérias envolvendo questões técnicas:

> At this point in the CPDP panel session, the moderator posed a somewhat provocative question to the audience: whether data controllers, in particular service providers, can mislead DPAs in ICT-related matters. From the three options: (1) frequently, (2) sometimes, (3) almost never, the majority of the audience voted for option 2. A panellist added that data controllers can mislead DPAs in such matters whenever they want but they do this only infrequently, because they are afraid of the risks.[475]

Neste ponto da pesquisa, os autores concluem que o nível de *expertise* das DPAs em questões técnicas é ao menos não satisfatório ou que necessita ser aprimorado:

---

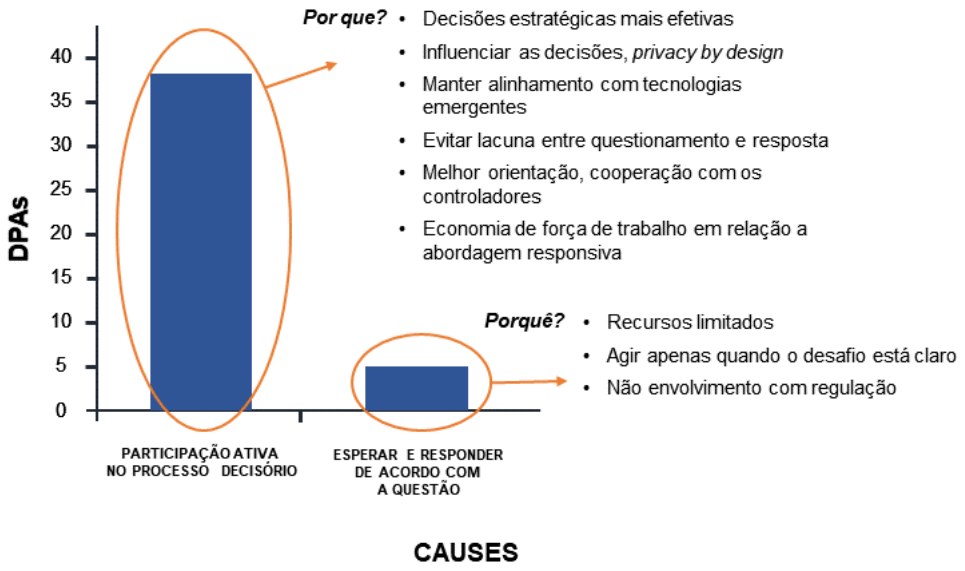
[475] RAAB, C.; SZEKELY, I. Data Protection Authorities and Information Technology. *Computer Law and Security Review*, 1 jul. 2017. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2994898. Acesso em: 14 maio 2021. p. 8-9.

> Both the analysis of the survey results and the opinions presented at the public discussion showed that the present level of expertise available in the DPAs' offices is either not satisfactory or at least needs to be further enhanced. With regard to rapid technological developments and their impact on the processing and use of personal data, it seems evident that even the mere preservation of the existing level of ICT expertise in an organisation requires continuous learning.[476]

Um terceiro resultado da pesquisa a ser destacado para os fins desta subseção diz respeito à abordagem das autoridades em relação à tomada de decisão em aspectos relacionados às TICs. Por tomada de decisão (*decision-making*) entende-se toda e qualquer atividade que exija um posicionamento da DPA, incluindo a elaboração de legislação, regulamentação ou edição de opiniões sobre certas operações de processamento de dados. De acordo com os autores, a pesquisa ofereceu duas opções aos entrevistados, perguntando se eles preferem (a) participar de forma proativa em tomadas de decisão estratégicas em questões de TIC relacionadas à privacidade, ou (b) se eles preferem esperar até que as implicações reais estejam claras para que possam responder de acordo com as circunstâncias. Conforme explicita a imagem abaixo, a quase totalidade das DPAs prefere uma abordagem proativa, tendo havido um pequeno número de autoridades que alegaram preferir uma abordagem responsiva.

[476] RAAB, C.; SZEKELY, I. Data Protection Authorities and Information Technology. *Computer Law and Security Review*, 1 jul. 2017. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2994898. Acesso em: 14 maio 2021. p. 9.

GRÁFICO 3 – Preferência das dpas por uma abordagem proativa ou responsiva



Fonte: RAAB; SXEKELY, 2017,p.11.

É interessante olhar para as causas que justificam a escolha da abordagem responsiva:

> The few DPAs that preferred reactive actions mentioned their *limited resources*, or the fact that their organization is not involved in the legislative process, however, one DPA clearly stated that the proper strategy is to act only when the challenge is clear.[477] (grifo nosso).

Para além do déficit de *expertise* técnica, as autoridades supervisoras europeias enfrentam ainda uma série de deficiências em termos de recursos humanos e financeiros e ausência de efetiva independência e autonomia funcional. Um estudo publicado pela European Union Agency for Fundamental Rights em 2010, ainda na vigência da Diretiva 95/46/CE, traça um diagnóstico preciso ao sumarizar o conjunto de desafios institucionais enfrentados pelas autoridades europeias à época:

> At a structural level, the lack of independence of several Data

[477] RAAB, C.; SZEKELY, I. Data Protection Authorities and Information Technology. *Computer Law and Security Review*, 1 jul. 2017. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2994898. Acesso em: 14 maio 2021. p. 11.

> Protection Authorities (DPAs) poses a major problem. In a number of Member States concerns are reported about the effectiveness and capability of the officers of Data Protection Authorities to perform their task with complete autonomy. At the functional level, understaffing and a lack of adequate financial resources among several Data Protection Authorities constitutes a major problem. At the operative level, a major problem is represented by the limited powers of several Data Protection Authorities. In certain Member States, they are not endowed with full powers to investigate, intervene in processing operations, offer legal advice and engage in legal proceedings.[478]

Embora a GDPR tenha buscado endereçar boa parte das deficiências presentes no contexto da Diretiva, o cenário após os dois primeiros anos de vigência do Regulamento parece não ser muito diferente, sendo caracterizado pela continuidade ou persistência de muitos dos desafios institucionais acima mencionados.

Em maio de 2020, a ONG AccessNow lançou o relatório "Two Years Under the EU GDPR: An Implementation Progress Report", no qual destaca ao menos três desafios relacionados ao *enforcement* do Regulamento: (i) uma aplicação ainda tímida das sanções econômicas, (ii) ausência de recursos humanos e financeiros suficientes e adequados ao desempenho das atividades institucionais e (iii) falência dos mecanismos de cooperação interinstitucional, em especial do *one-stop-shop*.[479]

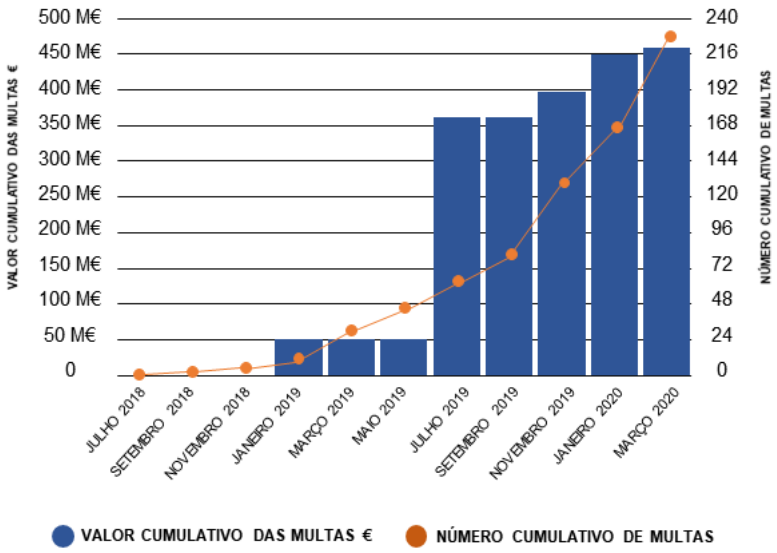
Até maio de 2019, as autoridades supervisoras aplicaram um total de 231 multas, sendo possível observar um significativo aumento na atividade de *enforcement* sancionatório entre julho de 2018 e maio de 2019 a partir do aumento do número de multas e do valor das sanções, conforme ilustrado no gráfico abaixo:

---

[478] EUROPEAN UNION AGENCY FOR FUNDAMENTAL RIGHTS. *Data Protection in the European Union*: the role of National Data Protection Authorities. Bruxelas: FRA, 2010. Disponível em: https://fra.europa.eu/sites/default/files/fra_uploads/815-Data-protection_en.pdf. Acesso em: 16 maio 2021. p. 6.

[479] MASSÉ, E. *Two Years Under the EU GDPR:* An Implementation Progress Report. [*S. l.*]: AccessNow, 2020. Disponível em: https://www.accessnow.org/cms/assets/uploads/2020/05/Two-Years-Under-GDPR.pdf. Acesso em: 16 maio 2021.

GRÁFICO 4 – Aplicação das sanções de multa nos primeiros 2 anos De vigência da GDPR



Fonte: MASSÉ, 2020,p.6.

Esse aumento no número de sanções econômicas aplicadas, contudo, deve ser analisado dentro de um contexto maior, considerado o volume de casos nos quais as autoridades são chamadas a se manifestar. Em relação ao exposto, o relatório aponta:

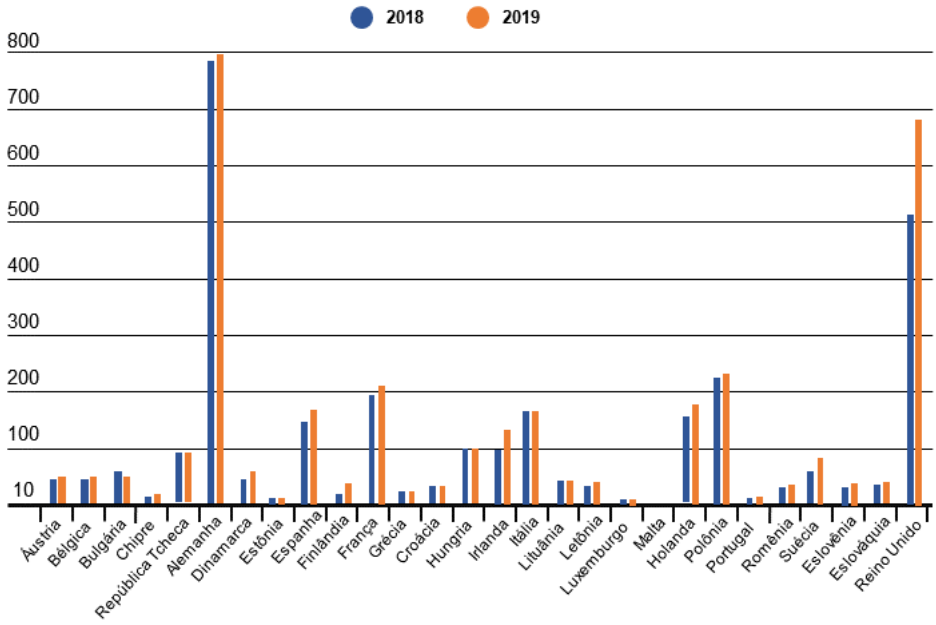
> While this growth is positive, *the total number of fines is still low compared to the 144,376 complaints that people had filed by May 2019*. While not every complaint will result in a fine and other sanctions are available as a remedy for data protection violations, a large number of complaints remain unaddressed. *DPAs are now facing a backlog of complaints*. In addition to watching any investigations they may launch on their own, we are waiting for DPAs to respond to these complaints to protect our rights.[480] (grifo nosso).

No relatório elaborado para o ano de 2018, a organização destacou que as DPAs, enquanto principais entidades responsáveis pela aplicação da GDPR, ocupariam um lugar central no sucesso ou insucesso do Regulamento, e que para garantir a sua efetiva implementação era necessário garantir-lhes maiores recursos

[480] MASSÉ, E. *Two Years Under the EU GDPR:* An Implementation Progress Report. [*S. l.*]: AccessNow, 2020. Disponível em: https://www.accessnow.org/cms/assets/uploads/2020/05/Two-Years-Under-GDPR.pdf. Acesso em: 16 maio 2021. p. 7.

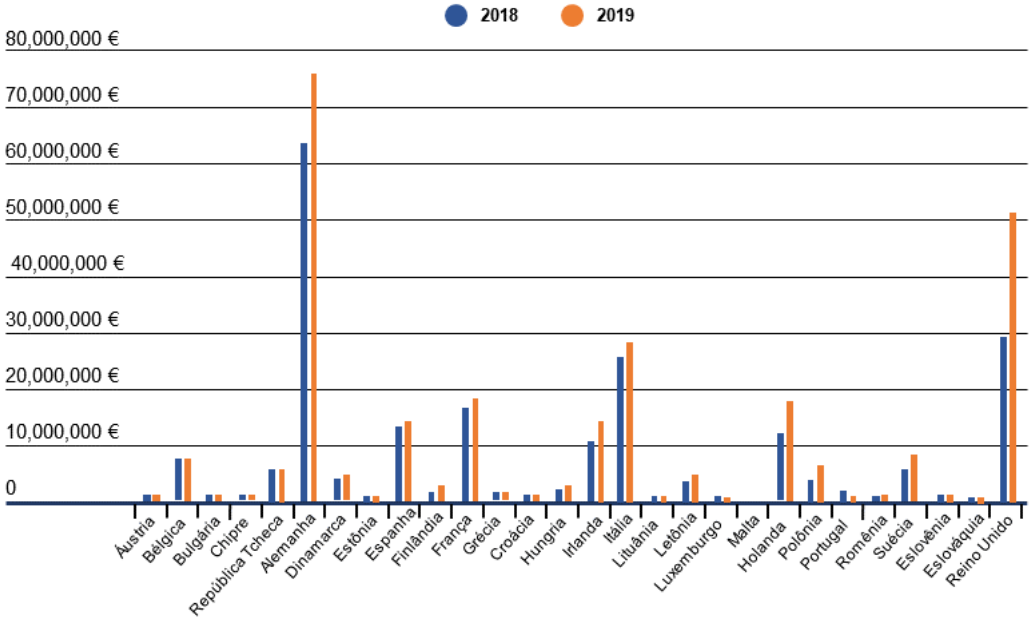
humanos e financeiros para o adequado desempenho de suas funções. Com base no relatório de 2020, a organização compara o aumento no número de funcionários e no orçamento de cada autoridade no período 2018-2019, destacando o tímido aumento de recursos humanos e financeiros, conforme ilustrado nos gráficos a seguir:

GRÁFICO 5 – Evolução do *staff* das DPAs entre 2018 e 2019



Fonte: MASSÉ, 2020,p.9.

GRÁFICO 6 – Evolução do orçamento das DPAs entre 2018 e 2019



Fonte: MASSÉ, 2020,p.10.

Como visto acima, houve um tímido aumento no número de funcionários de cada autoridade no período considerado. Conforme aponta o relatório, de acordo com informações compartilhadas por cada DPA com o European Data Protection Board, a expectativa para o ano de 2020 era de que o quadro permanecesse mais ou menos o mesmo.[481]

Quanto ao gráfico de evolução do orçamento, o relatório chama a atenção para as disparidades envolvendo os recursos financeiros disponíveis nas diferentes autoridades. A autoridade alemã desponta como a autoridade com maior número de recursos humanos e financeiros, mas é preciso considerar que esses recursos são distribuídos entre as 16 autoridades subnacionais e a autoridade federal alemã. Realizada esta consideração, o Reino Unido figura como o país com maior número de recursos financeiros e humanos voltados para o *enforcement* do Regulamento. O relatório chama atenção para o fato de que o Reino Unido não é propriamente o país com mais recursos, mas talvez o único país com um nível adequado de recursos para

[481] MASSÉ, E. *Two Years Under the EU GDPR:* An Implementation Progress Report. [*S. l.*]: AccessNow, 2020. Disponível em: https://www.accessnow.org/cms/assets/uploads/2020/05/Two-Years-Under-GDPR.pdf. Acesso em: 16 maio 2021. p. 9.

a execução das atividades de *enforcement*, pontuando a disparidade entre seu orçamento e quadro de funcionários com a de países com número de habitantes e posição econômica similar, como Itália e França, por exemplo:

> The United Kingdom is therefore the authority with the most staff and financial resources. It is also one of the few authorities to have in-house technological expertise, which is extremely important in conducting independent investigations. As the UK is set to exit the European Union, the network of European authorities working together within the EDPB will lose the support of this highly resourced authority. The UK's budget is double that of Italy and three times bigger than that of France, even though the three countries have roughly the same number of inhabitants and their economies are similar in size. And no, this does not mean that the UK has too big a budget; it may actually be the only authority with adequate resources.[482]

A situação da autoridade britânica parece ser a exceção no cenário europeu, uma vez que a maior parte das autoridades de *enforcement* não dispõe dos recursos necessários para a execução de suas atividades, o que coloca em risco a própria efetividade do Regulamento e os direitos dos titulares de dados:

> At the center of the success or failure of the GDPR are the Data Protection Authorities. If they do not enforce the law, we as individuals may never experience its benefits. For DPAs to function properly and address the large number of complaints that have been filed, the resources allocated to them must be increased. Politico Europe reported that many DPAs have been expressing their dissatisfaction with their current budget and resources. Out of 30 DPAs from all 27 EU countries, the United Kingdom, Norway, and Iceland, only nine said they were happy with their level of resourcing.[483]

Para além do déficit na garantia dos direitos positivados, a ausência de recursos adequados têm ainda efeitos reflexos no grau de confiança e no comportamento dos agentes regulados, que podem tender a ignorar o Regulamento por saberem que seu *enforcement* é deficitário ou inexistente, bem como limitar uma atuação mais rigorosa e incisiva das autoridades quando isto implicar maior emprego de recursos humanos e financeiros:

> EU states are failing and leaving their DPAs in a critical situation. If DPAs do not have adequate resources, we risk a return to the "business as usual" scenario that we experienced under the 95/46

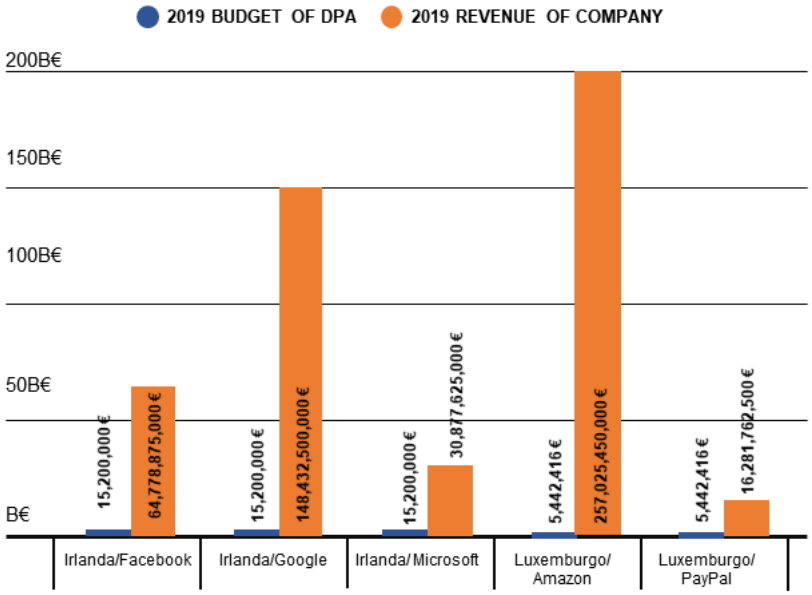
[482] MASSÉ, E. *Two Years Under the EU GDPR:* An Implementation Progress Report. [*S. l.*]: AccessNow, 2020. Disponível em: https://www.accessnow.org/cms/assets/uploads/2020/05/Two-Years-Under-GDPR.pdf. Acesso em: 16 maio 2021. p. 11.
[483] Ibidem, p. 9.

> Directive, when many companies ignored the law because enforcement was either slow or nonexistent. The inadequate budget provided to DPAs means that our rights may not be effectively protected. In fact, it may create a negative incentive for DPAs investigating large tech companies to agree on settlements that may be more favourable to the companies. The UK authority reached a settlement with Facebook following the Cambridge Analytica scandal and it is believed the authority opted for this avenue to limit the cost of lengthy proceedings. This is particularly striking as the UK authority is one of best-resourced 17 DPAs.[484]

Por fim, o déficit de recursos financeiros das autoridades de supervisão se torna ainda mais problemático quando analisado em face do enorme poder econômico dos *big players* submetidos ao seu escrutínio, que em muitas vezes ultrapassa o poder econômico das próprias autoridades e, em alguns casos, até mesmo o Produto Interno Bruto dos países ao qual suas atividades estão circunscritas. O gráfico abaixo ilustra essa disparidade ao comparar o orçamento de algumas DPAs com o faturamento de algumas das principais empresas sob escrutínio regulatório no cenário europeu:

[484] MASSÉ, E. *Two Years Under the EU GDPR:* An Implementation Progress Report. [*S. l.*]: AccessNow, 2020. Disponível em: https://www.accessnow.org/cms/assets/uploads/2020/05/Two-Years-Under-GDPR.pdf. Acesso em: 16 maio 2021. p. 9.

GRÁFICO 7 – Orçamento das DPAs vs Faturamento das companhias De tecnologia sua supervisão



Fonte: MASSÉ, 2020, p.11.

Essa disparidade entre entidade supervisora e agente regulado levanta sérias preocupações quanto ao poder dos agentes econômicos em tirar vantagem dessa condição em proveito próprio:

> Companies could leverage DPAs' lack of resources, using it to get around the application of the GDPR, or at least significantly delay its effect. The Irish authority has for instance indicated that "procedural queries" from companies are delaying what would be their first fines. In fact, Data Protection Authorities often lack the financial resources to enter into lengthy legal proceedings, which involve several layers of appeals, while companies are not so constrained. The graphic above illustrates the disparity of resources between Data Protection Authorities and the companies they are supposed to keep in check. Large tech companies have nearly endless financial resources in comparison to the restrictive budget allocated to Data Protection Authorities. In the case of Ireland, the revenue of some of these companies is even higher than the Gross Domestic Product of the country.[485]

Por fim, um terceiro aspecto que pode comprometer a efetiva garantia e

[485] MASSÉ, E. *Two Years Under the EU GDPR:* An Implementation Progress Report. [*S. l.*]: AccessNow, 2020. Disponível em: https://www.accessnow.org/cms/assets/uploads/2020/05/Two-Years-Under-GDPR.pdf. Acesso em: 16 maio 2021. p. 11-12.

implementação do direito à explicação diz respeito aos obstáculos para a construção de um ambiente regulatório uniforme no âmbito da UE. Conforme destacamos nos Capítulos 2 e 3, a Diretiva 95/46/CE, ao exigir que os Estados Membros incorporassem suas disposições ao direito interno por meio de leis nacionais, deu origem a uma fragmentação regulatória e à ausência de convergência no endereçamento de questões envolvendo proteção de dados pessoais no âmbito do bloco. A GDPR, enquanto um regulamento de aplicabilidade direta, aumentou o nível de convergência regulatória a nível regional, uma vez que os 27 Estados-Membros do bloco têm como referência um único instrumento legal. Em grande medida, a GDPR também buscou garantir uma maior harmonização e coesão no *enforcement* da legislação ao prever alguns mecanismos de cooperação institucional, como a cooperação entre diferentes DPAs, o *consistency mechanism* e o *one-stop-shop*. Orla Lynskey, contudo, argumenta que esses três mecanismos, inicialmente concebidos para pôr um fim ao problema da fragmentação das diversas leis nacionais de proteção de dados na UE, muito provavelmente não conseguirão atingir seus objetivos. Embora esta afirmação tenha sido feita ainda em 2015, com base no *draft* da GDPR, os diagnósticos mais recentes dos primeiros anos de aplicação da GDPR parecem corroborá-la, ao apontarem os desafios de implementação desses mecanismos.[486]

[486] "In December 2015, during the negotiations of the GDPR, the legal services of the Council which represents the EU states expressed concerns regarding the functioning of the one-stop-shop. They indicated that 'the lead authorities are a bad system if you want to protect citizens' fundamental rights', and noted further that while the system would be a one-stop-shop for companies, it would be "a three-stop-shop" for people, as we would have to deal with several authorities and courts to get a complaint resolved. At the time, these concerns were regarded as highly political, as they risked the extension of already lengthy negotiations of the law. Two years into the application of the law, these comments do unfortunately summarise the current situation. Several Data Protection Authorities are calling out the bottleneck of cross-border cases, as leading authorities are neither being transparent nor moving quickly enough to process complaints. Earlier this year, Ulrich Kelber, the head of Germany's federal Data Protection Authority, called the functioning of the current cross-border enforcement system 'unbearable'. A few months later, the Hamburg DPA called the one-stop-shop mechanism 'cumbersome, time consuming and ineffective'. One of the major hurdles for cooperation is, once again, budget and resources. Out of all EU countries, only five considered that they have enough resources to dedicate time to coordination tasks, including cross-border complaints. The five countries are the Czech Republic, Denmark, Hungary, the UK (which left the EU), and Luxembourg (which has yet to resolve any major case). All other countries report major issues. In the same survey, the Austrian Data Protection Authority said that they are not equipped to deal with some cases requiring cooperation with other authorities because '[a] lawyer of the authority is dealing with more than 100 cases (national and cross-border) simultaneously at an average'. Many countries, including Belgium, Lithuania, Poland, and Bulgaria indicated that the authority does not have staff allocated to this cooperation. The 17 German authorities collectively indicated that 'the current staffing is not found to be sufficient for the effective performance' of cooperation tasks. The Spanish authority notes that while they have more staff since 2018, 'the increase in staff is insufficient to meet the growth in workload derived from the GDPR'. The French authority indicates that it lacks the human resources to 'effectively contribute to all cooperation mechanisms'. Portugal's situation is the most alarming as 'there is only one person (almost entirely)

Além do enfraquecimento dos mecanismos de cooperação, Lynskey argumenta[487] que a consistência e a uniformidade do ambiente regulatório europeu são ainda dificultadas pelo amplo espaço conferido pelo Regulamento aos Estados-Membros para interpretarem o texto e aplicarem a legislação nacional:

> Following the publication of the Commission's Proposed Regulation, the EDPS noted that, although the Regulation makes significant advances towards the creation of a single applicable EU data protection law, its provisions leave more space than one might initially think for 'coexistence and interaction between EU law and national law'. The EDPS categorized the provisions of the proposed Regulation which interact with national law in four different ways: provisions which build upon national law, provisions which allow national law to build upon the Regulation, provisions which allow national law to depart from the Regulation, and provisions which allow national law to develop further the principles of the Regulation.[488]

Uma dessas aberturas conferidas pela GDPR diz respeito justamente ao direito à explicação. O art. 22(2)(b) assim dispõe: "O titular dos dados tem o direito de não ficar sujeito a nenhuma decisão tomada exclusivamente com base no tratamento automatizado, incluindo a definição de perfis, que produza efeitos na sua esfera jurídica ou que o afete significativamente de forma similar. 2. O nº 1 não se aplica se a decisão: [...] b) For autorizada pelo direito da União ou do Estado-Membro a que o responsável pelo tratamento estiver sujeito, e na qual estejam igualmente previstas medidas adequadas para salvaguardar os direitos e liberdades e os legítimos interesses do titular dos dados".

Em face deste dispositivo, os Estados-Membros têm adotado diferentes legislações com diferentes estratégias de implementação do art. 22(2)(b), que podem ser categorizadas de quatro formas. Primeiramente, tem-se a *abordagem negativa ou neutra.* Países com abordagem negativa não se valem da abertura dada pelo art. 22 para legislarem de modo diverso por meio de instrumentos internos. Ou simplesmente não legislam, ou se limitam a repetir o texto do art. 22 nas legislações nacionais. É

dedicated to that task." (MASSÉ, E. *Two Years Under the EU GDPR:* An Implementation Progress Report. [*S. l.*]: AccessNow, 2020. Disponível em: https://www.accessnow.org/cms/assets/uploads/2020/05/Two-Years-Under-GDPR.pdf. Acesso em: 16 maio 2021. p. 13).

487 LYNSKEY, O. *The Foundations of EU Data Protection Law*. Oxford: Oxford University Press, 2015. p. 70.

488 Ibidem, p. 71.

este o caso de 20 dos 27 países do bloco europeu.[489] Em segundo lugar, há a *abordagem setorial, por meio da qual os países implementam legislações nacionais regulamentando o art. 22(2)(b) somente em setores específicos.* É o caso da Alemanha, por exemplo, que adotou na sua lei de seguros uma permissão para a utilização de decisões automatizadas, permitida sempre que a decisão for favorável ao titular ou, quando não favorável, desde que sejam respeitadas determinadas regras vinculantes.[490] Terceiro, há a abordagem procedimental, que se vale da abertura dada pelo dispositivo para procedimentalizar o direito à explicação. Este é o caso do Reino Unido, que adotou uma lei nacional, o Data Protection Act 2018, na qual se procedimentaliza, de certa forma, um direito à explicação. Esse procedimento é dividido em três fases: (i) notificação da existência de decisão automatizada; (ii) requisição do titular e explicação e (iii) atendimento à requisição do titular. Adota ainda esta abordagem a Irlanda, que possui uma lei bastante similar à do Reino Unido.[491] Por fim, há a abordagem proativa, caracterizada por leis que trazem especificações extras ao art. 22(2)(b), sendo o caso da Loi 2018-493, de 20 de junho de 2018, da França.[492]

A partir deste cenário, é possível imaginar que haverá, no mínimo, diferentes entendimentos sobre a real extensão do direito à explicação, em razão das diferentes formas como o art. 22(2)(b) foi implementado nas legislações nacionais, o que pode gerar desafios institucionais de harmonização da interpretação em torno desse direito. Este e os demais problemas descritos ao longo desta seção, quais sejam, o déficit de *expertise* técnica das autoridades supervisoras e a ausência de recursos humanos e financeiros em volume adequado para garantir adequado exercício de suas funções, constituem alguns dos desafios institucionais de implementação do direito à explicação.

---

[489] SOUZA, C. A.; PERRONE, C.; MAGRANI, E. O direito à explicação entre a experiência europeia e a sua positivação na LGPD. *In*: MENDES, L. S.; DONEDA, D.; SARLET, I. W.; RODRIGUES JÚNIOR, O. L.; BIONI, B. (Orgs.). *Tratado de Proteção de Dados Pessoais*. Rio de Janeiro: Forense, 2021. p. 243-270. p. 254.

[490] Ibidem, p. 254-255.

[491] Ibidem, p. 256.

[492] SOUZA, C. A.; PERRONE, C.; MAGRANI, E. O direito à explicação entre a experiência europeia e a sua positivação na LGPD. *In*: MENDES, L. S.; DONEDA, D.; SARLET, I. W.; RODRIGUES JÚNIOR, O. L.; BIONI, B. (Orgs.). *Tratado de Proteção de Dados Pessoais*. Rio de Janeiro: Forense, 2021. p. 243-270.

4.3.2 ANPD: perspectivas e desafios

Como vimos na seção anterior, mesmo as autoridades de *enforcement* europeias, que contam com uma larga experiência jurídica e institucional acumulada, ao menos, desde a década de 1970, bem como garantias legais de independência e autonomia técnica, decisória e financeira solidamente estabelecidas, não estão livres de obstáculos e desafios institucionais no desempenho de suas funções. Neste sentido, é possível imaginar que tais desafios institucionais sejam ainda maiores no cenário brasileiro, tendo em vista o frágil e ainda inacabado contorno institucional da Autoridade Nacional de Proteção de Dados.

Assim como no cenário europeu, a existência de uma autoridade responsável pelo *enforcement* da nova regulação é peça fundamental para a garantia da eficácia da LGPD. Desde a primeira versão do Anteprojeto da Lei Geral de Proteção de Dados Pessoais, a previsão de uma lei acompanhada de uma autoridade de *enforcement* sempre esteve presente como modelo. O desenho institucional a ser assumido pela autoridade, contudo, foi e continua sendo um dos principais pontos de debate.[493]

Embora a composição desse tipo de órgão varie de país para país, há determinados atributos internacionalmente reconhecidos como sendo essenciais a uma autoridade de proteção de dados, sendo eles a garantia de independência e autonomia.

A independência é um atributo considerado intrínseco à própria razão de ser de uma autoridade de proteção de dados, havendo diversos meios de assegurá-la: garantindo-lhe independência técnica, financeira, estabelecendo critérios de nomeação dos integrantes, mandatos fixos, dentre outros. Conforme aponta Keller, a garantia de autoridades de proteção de dados independentes se justifica pelo fato de elas serem responsáveis por fiscalizar tanto o setor privado quanto o próprio governo, sendo a independência uma forma de garanti-las liberdade para fiscalizar os diferentes agentes sem sofrer influências ou pressões externas.[494] Para Doneda, a garantia de autoridades independentes encontra sua justificativa na necessidade de garantir respostas mais adequadas e céleres do que as que seriam oferecidas pela

---

[474] DONEDA, D. A Autoridade Nacional de Proteção de Dados e o Conselho Nacional de Proteção de Dados. *In*: MENDES, L. S.; DONEDA, D.; SARLET, I. W.; RODRIGUES JÚNIOR, O. L.; BIONI, B. (Coords.). *Tratado de Proteção de Dados Pessoais*. Rio de Janeiro: Forense, 2021. p. 459.

[494] KELLER, C. I. *Regulação nacional de serviços na Internet*: exceção, legitimidade e o papel do Estado. Rio de Janeiro: Lumen Juris, 2019. p. 243-244.

administração direta, que não necessariamente possui *expertise* técnica sobre questões envolvendo proteção de dados. Para o setor privado, a existência de autoridades independentes traz ainda benefícios como a "[...] uniformização da aplicação da lei em um mesmo território e em circunstâncias nas quais eventualmente tribunais ou reguladores setoriais tendessem a produzir soluções heterogêneas quanto à interpretação da legislação de proteção de dados."[495] A existência de um órgão independente também garante a

> [...] a centralização da matéria em uma autoridade evita o risco da fragmentação da interpretação da lei entre tribunais e mesmo outros órgãos administrativos com competências eventualmente concorrentes, garante a uniformidade dos direitos do cidadão e a segurança jurídica na aplicação da LGPD."[496]

Conforme assinala Doneda, para que se garanta a independência da autoridade, "suas atividades fiscalizatória, sancionatória e decisional não devem se subordinar hierarquicamente a outros órgãos.*"*[497]

Além dos atributos de autonomia e independência, há ainda uma outra característica necessária, que é a presença de um corpo técnico especializado (em assuntos jurídicos, regulatórios e técnicos):

> A autoridade é um elemento indispensável para garantir a adaptação da lei a novas circunstâncias sem que se abra mão da segurança jurídica, ao proporcionar orientação sobre a interpretação e aplicação da lei, ao elaborar normas e regulamentos sobre temas específicos como segurança da informação ou outras situações, sem que haja necessidade de alteração da lei. Ela pode ainda estabelecer parâmetros para a aplicação da lei conforme as características de cada setor ou mercado, objetivando ações que sejam mais eficazes para a proteção dos direitos do cidadão e garantindo a proporcionalidade na sua aplicação – considerando, por exemplo, o seu impacto em pequenas e médias empresas. *Para tanto, contar com pessoal técnico especializado é um elemento de primeira importância*.[498] (grifo nosso).

De acordo com Doneda, o preenchimento desses atributos é não apenas

---

[495] DONEDA, D. A Autoridade Nacional de Proteção de Dados e o Conselho Nacional de Proteção de Dados. *In*: MENDES, L. S.; DONEDA, D.; SARLET, I. W.; RODRIGUES JÚNIOR, O. L.; BIONI, B. (Coords.). *Tratado de Proteção de Dados Pessoais*. Rio de Janeiro: Forense, 2021. p. 465.
[496] Ibidem.
[497] Ibidem, p. 466.
[498] DONEDA, D. A Autoridade Nacional de Proteção de Dados e o Conselho Nacional de Proteção de Dados. *In*: MENDES, L. S.; DONEDA, D.; SARLET, I. W.; RODRIGUES JÚNIOR, O. L.; BIONI, B. (Coords.). *Tratado de Proteção de Dados Pessoais*. Rio de Janeiro: Forense, 2021. p. 466.

necessário para o bom funcionamento da autoridade e para o efetivo *enforcement* da legislação, mas também condição para que o Brasil usufrua das vantagens políticas e econômicas decorrentes da LGPD, como a acessão à OCDE, a obtenção de nível de adequação junto à UE (cuja existência de uma autoridade independente é condição indispensável), dentre outros aspectos.[499]

Mas em que medida o atual desenho institucional da ANPD contempla essas características? E como sua presença ou ausência pode impactar a garantia e efetiva implementação do direito à explicação no cenário brasileiro?

A previsão de criação de um órgão nos moldes de uma autoridade independente estava presente já em 2018, quando a Comissão Especial criada para analisar o PL nº 5276/2016 emitiu relatório contendo proposta de criação de um órgão na forma de uma autarquia federal em regime especial, dotado de independência e autonomia técnica, decisória e financeira, que foi encaminhada ao Senado Federal na forma do PLC nº 53/2018. Esse modelo de estrutura administrativa foi aprovado por ambas as casas do Congresso Nacional, mas sofreu veto presidencial sob o fundamento de vício de competência.[500] Em 14 de agosto de 2018, a LGPD foi aprovada com a criação da ANPD vetada, mantendo-se no texto as várias menções à autoridade, comprometendo a lei em vários pontos. Em dezembro do mesmo ano, contudo, a Medida Provisória nº 869/2018, convertida na Lei nº 13.853/2019, finalmente cria a Autoridade Nacional de Proteção de Dados, porém com um desenho institucional bastante diferente daquele que havia sido discutido e aprovado no Congresso: a ANPD é criada não nos moldes de uma autarquia federal especializada, integrante da administração pública indireta, mas como um órgão da administração pública direta vinculado à Presidência da República.[501] Não obstante, o texto estabelece que a natureza jurídica da ANPD possui caráter provisório, podendo sofrer uma revisão de seu desenho institucional dentro do prazo de dois anos (art. 55-A, §§ 1º e 2º). Cabe ainda destacar que a ANPD já nasce sob fortes cortes e restrições orçamentárias. A autoridade depende do orçamento da Presidência da República e, como não há a possibilidade de criação de novos cargos, todo o preenchimento de

[499] DONEDA, D. A Autoridade Nacional de Proteção de Dados e o Conselho Nacional de Proteção de Dados. *In*: MENDES, L. S.; DONEDA, D.; SARLET, I. W.; RODRIGUES JÚNIOR, O. L.; BIONI, B. (Coords.). *Tratado de Proteção de Dados Pessoais*. Rio de Janeiro: Forense, 2021.
[500] Ibidem, p. 467.
[501] DONEDA, D. A Autoridade Nacional de Proteção de Dados e o Conselho Nacional de Proteção de Dados. *In*: MENDES, L. S.; DONEDA, D.; SARLET, I. W.; RODRIGUES JÚNIOR, O. L.; BIONI, B. (Coords.). *Tratado de Proteção de Dados Pessoais*. Rio de Janeiro: Forense, 2021. p. 467.

seus quadros técnicos tem se dado por meio de requisição. Ademais, dentre os 9 pontos da LGPD vetados pelo presidente Bolsonaro, estava justamente a possibilidade de a ANPD cobrar taxas pelos serviços prestados.[502] Como esses déficits institucionais e a ausência de recursos humanos e financeiros adequados irão impactar o *enforcement* da LGPD e a garantia do direito à explicação é uma questão ainda em aberto, cabendo observar como a autoridade se comportará e reagirá em face desse cenário. Com base na experiência europeia, contudo, que conta com um arranjo institucional muito mais robusto e consolidado, mas que, ainda assim, enfrenta graves entraves às atividades de *enforcement*, é possível antever um cenário de inúmeros desafios nos primeiros anos de atuação da ANPD.

Além do desenho institucional da ANPD, os gargalos institucionais para a garantia do direito à explicação no Brasil podem ser ainda entendidos a partir dos desafios de coordenação intergovernamental para a implementação da lei. Conforme aponta Wimmer, a LGPD é uma lei de caráter transversal, que produz efeitos horizontais sobre todos os setores econômicos e sobre quase todos os campos de atuação do Poder Público, em seus diferentes níveis de atuação. De acordo com a autora, essa incidência transversal traz consigo desafios relacionados ao cruzamento de competências de outros órgãos e entidades da administração pública, em seus diferentes níveis de atuação, seja direta ou indireta: "Tal complexidade suscita importantes desafios hermenêuticos e impõe às diversas instâncias do Poder Executivo, em nome da segurança jurídica, um gigantesco desafio de coordenação."[503]

Wimmer aponta ainda que o advento de uma norma com características de transversalidade como a LGPD "[...] acarreta inúmeros desafios hermenêuticos

---

[502] "A lei também teve um veto sobre a cobrança de emolumentos pelos serviços prestados como fonte de eventual de recursos financeiros para a Autoridade Nacional de Proteção de Dados. A incapacidade de cobranças por serviços da ANPD retira a possibilidade de a entidade ter, a médio e longo prazo, receita e estrutura para corresponder às demandas e necessidades que surgirão no seu funcionamento." (COALIZÃO DIREITOS NA REDE. Coalizão Direitos na Rede repudia os 9 vetos de Bolsonaro à lei que cria a Autoridade Nacional de Proteção de Dados. *Medium*, 9 jul. 2019. Disponível em: https://cdr-br.medium.com/coaliz%C3%A3o-direitos-na-rede-repudia-os-9-vetos-de-bolsonaro-%C3%A0-lei-que-cria-a-autoridade-nacional-de-ee536f6baeb. Acesso em: 28 abr. 2021). Cf., também: BRASIL. Câmara dos Deputados. Sancionada, com nove vetos, lei que cria Autoridade Nacional de Proteção de Dados. *Câmara dos Deputados*, 9 jul. 2019. Disponível em: https://www.camara.leg.br/noticias/561908-SANCIONADA,-COM-NOVE-VETOS,-LEI-QUE-CRIA-AUTORIDADE-NACIONAL-DE-PROTECAO-DE-DADOS. Acesso em: 28 abr. 2021.

[503] WIMMER, M. Os desafios do *enforcement* na LGPD: fiscalização, aplicação de sanções administrativas e coordenação intergovernamental. *In:* MENDES, L. S.; DONEDA, D.; SARLET, I. W.; RODRIGUES JR., O. L. (Coords.). *Tratado de Proteção de Dados Pessoais*. Rio de Janeiro: Forense, 2021. p. 377.

decorrentes de sua interação com outras normas gerais e especiais existentes.” De acordo com a autora, essas dificuldades hermenêuticas se manifestam também “[...] no que tange ao arranjo institucional e à divisão de competências na Administração Pública.”[504] Nesse sentido, aponta que:

> A variedade de normas provenientes de diferentes microssistemas normativos e a diversidade de estruturas decisórias a regular uma mesma matéria fática geram a necessidade de arranjos institucionais e de ferramentas interpretativas capazes de resolver colisões entre os múltiplos centros de poder que caracterizam a Administração Pública contemporânea.[505]

O desafio de cooperação intergovernamental, articulação de competências e diálogo entre diferentes legislações setoriais definido pela autora se apresenta de forma bastante clara quando analisamos a garantia e implementação do direito à explicação. Como analisamos no Capítulo 3 do presente trabalho, não apenas a LGPD, como também o Código de Defesa do Consumidor e a Lei do Cadastro Positivo, trazem importantes normas versando sobre obrigações de transparência algorítmica e o direito à explicação no contexto das relações consumeristas. Como vimos, os dispositivos presentes nessas duas legislações têm sido amplamente aplicados pelo Poder Judiciário, dando origem a uma rica jurisprudência, e atraído ainda a competência dos órgãos integrantes do Sistema Nacional de Defesa do Consumidor. Como esse diálogo interinstitucional será estabelecido ainda é uma questão a ser esclarecida a curto e médio prazo, especialmente em relação à articulação da ANPD com o Poder Judiciário e sua produção jurisprudência sobre o tema. Quanto ao diálogo institucional estabelecido com os órgãos de defesa do consumidor, contudo, a ANPD parece já estar apostando numa lógica de cooperação com a Senacon e outras instâncias, conforme se depreende da postura assumida diante dos primeiros desafios de *enforcement* que lhe foram apresentados até então.[506]

---

[504] Ibidem, p. 380-381.
[505] Ibidem, p. 381.
[506] AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS. ANPD e Senacon assinam acordo de cooperação técnica. *ANPD*, 22 mar. 2021. Disponível em: https://www.gov.br/anpd/pt-br/assuntos/noticias/anpd-e-senacon-assinam-acordo-de-cooperacao-tecnica. Acesso em: 16 maio 2021; AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS. Cade, MPF, ANDP e Senacon recomendam que WhatsApp adie entrada em vigor da nova política de privacidade. *ANPD*, 7 maio 2021. Disponível em: https://www.gov.br/anpd/pt-br/assuntos/noticias/cade-mpf-anpd-e-senacon-recomendam-que-whatsapp-adie-entrada-em-vigor-da-nova-politica-de-privacidade. Acesso em: 21 jun. 2021.

Neste contexto de múltiplos órgãos com competências sobrepostas voltadas à garantia e implementação do direito à explicação, é crucial que a ANPD assuma seu protagonismo na interpretação e execução da LGPD[507], nos termos dos arts. 55-J, XX e 55-K, P.U., garantindo maior segurança jurídica aos indivíduos e organizações e buscando garantir que o direito à explicação não seja implementado de forma fragmentada, divergente e descentralizada. Neste contexto, considerando que a LGPD não faz menção expressa a um direito à explicação, é crucial que a autoridade se posicione acerca da existência, natureza jurídica e extensão deste direito, bem como sobre a interpretação que deve ser dada ao direito à revisão de decisão automatizada (se ampliativa, requerendo a revisão por pessoa natural, ou literal, nos termos da atual redação da lei). Tendo ainda em vista a forte proteção conferida ao segredo de negócio na LGPD, conforme discutido no item 4.1, caberá à ANPD esclarecer em que medida o direito à explicação e obrigações de transparência podem ser limitados pela proteção dada pelos direitos de propriedade intelectual sobre algoritmos.

Por fim, conforme discutimos no Capítulo 3, Mulholland e Frahjof fazem uma importante ressalva quanto à competência da Autoridade Nacional de Proteção de Dados para realizar auditorias em sistemas algorítmicos na hipótese de o controlador se recusar, alegando segredo comercial ou industrial, a fornecer informações claras e adequadas a respeito dos critérios e procedimentos empregados na decisão automatizada. A redação do art. 20, § 2º, expressamente prevê que "[...] a autoridade nacional *poderá* realizar auditoria" (grifo nosso), o que, para as autoras, denota a existência de uma margem de discricionariedade da ANPD para exercer ou não sua competência fiscalizatória. Tal como está posta, esta redação pode dar margem a comportamentos indesejados e oportunistas por parte do ente regulado:

> A segunda reconhece, à primeira vista, a discricionariedade da autoridade nacional para realizar a auditoria apenas quando o controlador se negar a fornecer as informações elencadas no parágrafo primeiro. A existência desta condição pode dar margem para que o controlador se negue a exercer a explicação com base em uma simples alegação de que seu código estaria protegido pelo segredo comercial ou industrial, pois sabe que a atuação da autoridade em auditar seu algoritmo será optativa. Por sua vez, caso a explicação concedida ao titular não seja suficientemente clara –

[507] CENTRE FOR INFORMATION POLICY LEADERSHIP. O papel da Autoridade Nacional de Proteção de Dados (ANPD) conforme a nova Lei Geral de Proteção de Dados Pessoais (LGPD). *CIPL*, 17 abr. 2020. p. 4.

> impedindo que a pessoa seja capaz de compreender se houve ou não um tratamento discriminatório, ou o motivo pelo qual o algoritmo decidiu de uma maneira e não de outra –, parece que o titular de dados terá menos garantias do que se o controlador de dados tivesse meramente alegado a proteção do sigilo do seu algoritmo.[508]

A partir desta análise do art. 20, § 2º, da LGPD, parece haver uma brecha a ser explorada pelos agentes de tratamento que pode comprometer significativamente a garantia e implementação do direito à explicação na LGPD.

4.3.3 Tribunais e outras instâncias administrativas

No cenário brasileiro, um segundo desafio institucional relacionado ao *enforcement* do direito à explicação reside na ampla abertura à tutela coletiva da proteção de dados pessoais conferida pela LGPD, o que dá origem a um complexo arranjo institucional em torno da aplicação desse direito, sobretudo quando consideradas as relações consumeristas. Ao mesmo tempo em que essa abertura à tutela coletiva pode ser positiva, ao ampliar, por exemplo, a tutelabilidade dos direitos dos titulares de proteção de dados pessoais, ela traz também alguns desafios relacionados à consistência da aplicação e harmonização da interpretação deste direito num contexto marcado pela coexistência de múltiplos atores competentes para analisar conflitos envolvendo obrigações de transparência e o direito à explicação.

Esses desafios institucionais podem ser mais bem compreendidos a partir das lentes propostas por Zanatta, que aponta estar em curso, globalmente, mas no Brasil sobretudo, um processo de "coletivização da proteção de dados pessoais". De acordo com o autor, é possível verificar a ocorrência de uma mudança estrutural nas leis de proteção de dados pessoais ao redor do mundo, que começam a migrar de um paradigma liberal clássico, pautado em direitos e mecanismos de tutela e reparação meramente individuais, para um paradigma de proteção coletiva, sendo exemplos claros desse movimento a própria GDPR e a LGPD. Conforme assinala, as transformações tecnológicas ocorridas nas últimas décadas e a própria configuração das relações envolvendo o tratamento de dados colocam em xeque a capacidade do indivíduo, isoladamente, autodeterminar-se em relação ao fluxo de seus dados, do

[508] MULHOLLAND, C.; FRAJHOF, I. Z. Inteligência Artificial e a Lei Geral 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. (Coords.). *Inteligência Artificial e Direito:* Ética, Regulação e Responsabilidade. 2. ed. São Paulo: Revista dos Tribunais, 2020.

que emerge a necessidade de se conceber a proteção de dados pessoais como uma agenda de direitos coletivos, não mais restrita ao paradigma individual, processo que o autor denomina de “coletivização da proteção de dados pessoais”.[509]

A coletivização da proteção de dados pessoais, de acordo com o autor, pode ser descrita a partir de quatro elementos essenciais. Primeiramente, observa-se um reforço a uma linguagem de direitos difusos e direitos coletivos, fazendo com que os casos sejam avaliados a partir de uma lógica de dano à sociedade ou de violação aos valores da sociedade. Segundo, observa-se uma mudança na forma de proteção dos direitos ou uma ampliação de sua tutelabilidade, na medida em que não só os indivíduos, mas também organizações da sociedade civil, passam a ser legitimados para defender os direitos de proteção de dados pessoais. Terceiro, verifica-se uma ampliação das obrigações de proteção do “ambiente informacional”, numa perspectiva preventiva e baseada em riscos. Por fim, o quarto elemento da coletivização pode ser entendido como

> [...] a redefinição das estruturas administrativas de defesa do consumidor, que passam a encarar a proteção de dados pessoais como um problema coletivo de “defesa do consumidor”, sendo exemplos desse movimento a atuação da Federal Trade Commission, nos EUA, e da Secretaria Nacional do Consumidor (Senacon), no Brasil.[510]

Além desses quatro elementos caracterizadores do processo de coletivização da proteção de dados pessoais quando globalmente considerado, há ainda um quinto elemento a ser levado em consideração no cenário brasileiro. Conforme aponta o autor,

> Além dessas quatro características gerais, o Brasil apresenta um quinto elemento qualificador da coletivização da proteção de dados pessoais, distinto de outras jurisdições caracterizado pela intensa atuação repressiva do Ministério Público (MP) na construção de ações civis públicas.[511]

---

[509] ZANATTA, R. A. F. A tutela coletiva na proteção de dados pessoais. *Revista do Advogado*, n. 144, nov. 2019. Disponível em: https://aplicacao.aasp.org.br/aasp/servicos/revista_advogado/paginaveis/144/200/index.html. Acesso em: 20 maio 2021. p. 202.

[510] ZANATTA, R. A. F. A tutela coletiva na proteção de dados pessoais. *Revista do Advogado*, n. 144, nov. 2019. Disponível em: https://aplicacao.aasp.org.br/aasp/servicos/revista_advogado/paginaveis/144/200/index.html. Acesso em: 20 maio 2021. p. 203-204.

[511] Ibidem, p. 204.

Nesse ponto, o autor destaca a atuação repressiva do Ministério Público do Distrito Federal e Territórios (MPDFT), por meio de sua Comissão Especial de Proteção de Dados Pessoais, que instaurou uma série de inquéritos contra empresas que haviam supostamente violado a LGPD, antes mesmo de sua entrada em vigência. Em um desses casos, o MPDFT teve uma atuação bastante problemática ao exigir de uma empresa a elaboração de relatório de impacto à proteção de dados pessoais, competência que é exclusiva da ANPD, nos termos do art. 38 da LGPD. Essa atuação repressiva do Ministério Público pode ser constatada ainda pelas inúmeras ações coletivas de grande repercussão em proteção de dados pessoais pautadas entre 2016 e 2018 pelos diversos braços do órgão, dentre as quais se destacam as seguintes: *Ministério Público Federal do Piauí vs. Google Brasil* (2016), *Ministério Público do Rio de Janeiro vs. Fetranspor* (2017), *Ministério Público Federal de São Paulo vs. Microsoft* (2018), *Ministério Público do Rio de Janeiro vs. Decolar* (2018), *Ministério Público do Distrito Federal e Territórios vs. Banco Inter* (2018) e *Ministério Público do Distrito Federal e Territórios vs. Telefônica* (2018).[512] A partir dessa forte atuação e protagonismo do Ministério Público, o autor argumenta que "Independentemente da avaliação sobre os excessos e eventuais ilegalidades da atuação do MP, o fato é que a coletivização da proteção de dados pessoais no Brasil possui a tendência de ser fortemente marcada pela atuação do MP."[513]

De modo geral, essa forte atuação do Ministério Público e a tendência de coletivização das questões envolvendo proteção de dados pessoais no Brasil pode ser explicada pela forte tradição de direitos coletivos e difusos existente no cenário brasileiro, que desde a década de 1980 tem sido bastante permeável ao debate teórico sobre acesso à justiça e tutela coletiva, o que se reflete em diversos dos instrumentos e mecanismos jurídicos adotados, como a Lei de Ação Civil Pública e o Código de Defesa do Consumidor, que diminuem as barreiras de acesso ao judiciário e ampliam a tutelabilidade dos direitos no Brasil.[514]

Em grande medida, essa tradição de direitos coletivos e difusos foi refletida no

---

[512] Para uma análise mais detalhada dos casos, cf. ZANATTA, R. A. F. Tutela coletiva e coletivização da proteção de dados pessoais. *In:* PALHARES, F. (Org.). *Temas Atuais de Proteção de Dados*. São Paulo: Thomson Reuters, 2020. p. 357-359.

[513] ZANATTA, R. A. F. A tutela coletiva na proteção de dados pessoais. *Revista do Advogado*, n. 144, nov. 2019. Disponível em: https://aplicacao.aasp.org.br/aasp/servicos/revista_advogado/paginaveis/144/200/index.html. Acesso em: 20 maio 2021. p. 204.

[514] Idem. Tutela coletiva e coletivização da proteção de dados pessoais. *In:* PALHARES, F. (Org.). *Temas Atuais de Proteção de Dados*. São Paulo: Thomson Reuters, 2020. p. 357-359.

texto da LGPD. Nas palavras do autor: “Não é sem razão que a LGPD possui uma profunda conexão com o CDC e o sistema de tutela coletiva. A LGPD absorveu parte da tradição de tutela coletiva no Brasil, abrindo espaço para que a proteção dos direitos assegurados na legislação seja feita de forma coletiva, ao lado das múltiplas formas de proteção individual dos direitos.”[515] Há uma série de dispositivos da LGPD que apontam para essa abertura à tutela coletiva, dentre os quais pode se destacar o art. 18, § 8º, que prevê que os titulares também poderão peticionar em relação aos seus direitos junto aos organismos de defesa do consumidor, o art. 22, que dispõe que a defesa dos interesses e dos direitos dos titulares de dados poderá ser exercida em juízo, individual ou coletivamente, na forma do disposto na legislação pertinente, acerca dos instrumentos de tutela individual e coletiva, e o art. 42, § 3º, que prevê que as ações de reparação por danos coletivos que tenham por objeto a responsabilização dos agentes de tratamento podem ser exercidas coletivamente em juízo. Essas disposições trazem uma série de implicações institucionais, uma vez que atraem para o *enforcement* dos direitos dos titulares, dentre os quais o direito à explicação, uma série de atores, com extensa e acumulada experiência de atuação na proteção de dados pessoais e que deverão atuar paralelamente à ANPD, como o Judiciário em suas diferentes instâncias e os órgãos integrantes do Sistema Nacional de Defesa do Consumidor, como os Procons e a Secretaria Nacional do Consumidor. Essa multiplicidade de atores com competências concorrentes ou sobrepostas em relação ao *enforcement* dos direitos dos titulares no cenário brasileiro é bem destacada pelo autor:

> A legislação é clara ao afirmar que o direito de peticionamento pode ser exercido “perante os organismos de defesa do consumidor” (art. 18, §8º), *reforçando a estrutura das centenas de Procons criados desde a década de 1970 no país. Mas não só. Essa norma atrai para o polo de aplicação desses direitos toda a estrutura dos Procons, Defensorias Públicas, ONGs e Ministérios Públicos (o que é chamado “Sistema Nacional de Defesa do Consumidor”)*, na medida em que a LGPD também afirma que “a defesa dos interesses e dos direitos dos titulares de dados poderá ser exercida em juízo, individual ou coletivamente”, “acerca dos instrumentos de tutela individual e coletiva” (art. 22).[516] (grifo nosso)
>
> De modo muito claro, *a LGPD estabelece que o judiciário poderá ser instado pelos legitimados, a defender os interesses e os direitos dos*

[515] ZANATTA, op. cit., 2019, p. 205.
[516] ZANATTA, R. A. F. Tutela coletiva e coletivização da proteção de dados pessoais. *In:* PALHARES, F. (Org.). *Temas Atuais de Proteção de Dados*. São Paulo: Thomson Reuters, 2020. p. 362.

> *titulares de dados, individual ou coletivamente*, na forma do disposto na legislação pertinente, acerca dos instrumentos de tutela individual e coletiva (art. 22).[517] (grifo nosso)
>
> Além disso, a LGPD apresenta-se como uma legislação de direitos difusos pela clara interação entre LGPD e Código de Defesa do Consumidor, abrindo caminho para uma coordenação institucional entre Secretaria Nacional do Consumidor e Autoridade Nacional de Proteção de Dados Pessoais. Além disso, a LGPD é bastante única, em nível mundial, ao promover uma interação explícita com o Sistema Nacional de Defesa do Consumidor no exercício dos direitos dos titulares." [...] "*De fato, como já sustentado, 'o enforcement se dará não só pela Autoridade Nacional de Proteção de Dados, mas também por meio desse arranjo de atores institucionais dedicados à tutela coletiva da proteção de dados pessoais*.[518] (grifo nosso).

Conforme afirmado no início desta seção, essa abertura à tutela coletiva pode ser positiva, ao ampliar a tutelabilidade do direito à explicação, mas traz também alguns desafios institucionais relacionados à consistência da aplicação e harmonização da interpretação deste direito.

Nos termos do art. 55-K, P.U., da LGPD, a ANPD deverá articular sua atuação com outros órgãos e entidades com competências sancionatórias e normativas afetas ao tema de proteção de dados pessoais e será o órgão central de interpretação da lei e do estabelecimento de normas e diretrizes para a sua implementação.[519] Embora a LGPD ocupe lugar de centralidade na interpretação da lei, sendo ainda responsável por sua regulamentação, é necessário pontuar que a ANPD, ainda em estágio inicial de sua atuação, ainda não se afirmou como um órgão central de interpretação, ao menos não em sua interface com o Poder Judiciário. Entre a promulgação da lei, o início de sua vigência e a efetiva implementação da ANPD, houve um significativo lapso temporal no qual predominou certa incerteza e insegurança jurídica entre os agentes regulados e os diferentes atores responsáveis pelo *enforcement* da lei, o que

---

[517] ZANATTA, R. A. F. Tutela coletiva e coletivização da proteção de dados pessoais. *In:* PALHARES, F. (Org.). *Temas Atuais de Proteção de Dados*. São Paulo: Thomson Reuters, 2020. p. 362.
[518] Ibidem, p. 363.
[519] As recomendações recentemente emitidas ao WhatsApp Inc. em conjunto pela Autoridade Nacional de Proteção de Dados, pelo Conselho Administrativo de Defesa Econômica, pelo Ministério Público Federal e pela Secretaria Nacional do Consumidor, recomendando que a empresa adiasse a entrada em vigência de sua nova política de privacidade e estabelecendo uma série de recomendações regulatórias, é um claro exemplo de exercício desse dever de articulação na prática, abordagem que deverá ser recorrente na atuação da ANPD. Cf. AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS. Cade, MPF, ANDP e Senacon recomendam que WhatsApp adie entrada em vigor da nova política de privacidade. *ANPD*, 7 maio 2021. Disponível em: https://www.gov.br/anpd/pt-br/assuntos/noticias/cade-mpf-anpd-e-senacon-recomendam-que-whatsapp-adie-entrada-em-vigor-da-nova-politica-de-privacidade. Acesso em: 21 jun. 2021.

comprometeu a compreensão de sua exata extensão e correta interpretação, cenário que continua presente e que perdurará por mais algum tempo, enquanto a autoridade não regulamentar as principais questões em aberto na legislação. A regulamentação sobre aspectos dos direitos dos titulares, por exemplo, está alocada na última fase da agenda regulatória para o biênio 2021-2022 recentemente divulgada pela ANPD, com previsão de início apenas para o primeiro semestre de 2022.[520] Não obstante esse cenário de ausência de uniformidade na aplicação da lei, fato é que o Poder Judiciário já tem começado a atuar de forma bastante marcante no *enforcement* da LGPD, o que pode ser considerado positivo, de um lado, ao se considerar que a ANPD ainda não iniciou seu *enforcement* sancionatório, mas também negativo quando se consideram os riscos de aplicação e interpretação descentralizada e assistemática da LGPD. Essas duas visões são bem pontuadas por Zanatta, que parece se filiar a um diagnóstico de matriz mais otimista:

> A LGPD garante ampla "tutelabilidade" aos interesses difusos de proteção de dados pessoais, abrindo a possibilidade de instrumentos jurídicos de matriz coletiva, como as Ações Civis Públicas. Essa é uma característica que tende a tornar a regulação muito mais complexa do que em outros países, tendo à vista a inação por parte do Poder Executivo em constituição da Autoridade Nacional de Proteção de Dados Pessoais, o que impede um deslocamento dos conflitos para a arena da tutela administrativa. [...] *Esse "não deslocamento", que tende a inflar ainda mais o Poder Judiciário inicialmente, pode ser visto de forma positiva por uma perspectiva de construção de cultura de proteção de dados pessoais, de capacitação de agentes e de democracia de alta energia baseada no conflito e na contestação. Esta visão se contrapõe a uma narrativa de que os litígios simplesmente gerariam insegurança jurídica e consequências negativas*.[521]

Danilo Doneda, por sua vez, ao discorrer sobre a importância de uma autoridade independente responsável pelo *enforcement* da LGPD, argumenta que ela se faz necessária, dentre outros motivos, para garantir "[...] a uniformização da aplicação da lei em um mesmo território e em circunstâncias nas quais eventualmente tribunais ou reguladores setoriais tendessem a produzir soluções heterogêneas quanto à interpretação da legislação de proteção de dados" e que

---

[520] BRASIL. Autoridade Nacional de Proteção de Dados. *Portaria nº 11, de 27 de janeiro de 2021*. Torna pública a agenda regulatória para o biênio 2021-2022. Diário Oficial da União, Brasília-DF, Seção 1, edição 19, publicado em 28 jan. 2021, p. 3. Disponível em: https://www.in.gov.br/en/web/dou/-/portaria-n-11-de-27-de-janeiro-de-2021-301143313. Acesso em: 21 maio 2021.

[521] ZANATTA, R. A. F. Tutela coletiva e coletivização da proteção de dados pessoais. *In:* PALHARES, F. (Org.). *Temas Atuais de Proteção de Dados*. São Paulo: Thomson Reuters, 2020. p. 370-371.

> [...] a centralização da matéria em uma autoridade evita o risco da fragmentação da interpretação da lei entre tribunais e mesmo outros órgãos administrativos com competências eventualmente concorrentes, garante a uniformidade dos direitos do cidadão e a segurança jurídica na aplicação da LGPD.[522]

Não obstante os ganhos que uma aplicação descentralizada e distribuída entre diferentes atores possa trazer para a conformação de uma cultura jurídica de proteção de dados pessoais no Brasil, conforme aponta por Zanatta, a ausência de critérios claros de interpretação e aplicação do direito à explicação, aqui especialmente considerado, que pode ser entendido, como vimos anteriormente, como um dos direitos mais controversos dentre aqueles constitutivos do plexo de direitos dos titulares, justamente em razão da existência de dúvidas e controvérsias sobre sua própria existência e escopo, parece ser problemática.

A abertura a uma tutela coletiva da LGPD e a possibilidade de *enforcement* descentralizado parece suscitar desafios ainda maiores ao direito à explicação, um direito que não apenas tende a ser, mas já tem sido efetivamente bastante explorado pelos tribunais e pelos órgãos de defesa do consumidor, justamente pela sua íntima relação com questões consumeristas, como os cadastros positivos e negativos de crédito e o emprego da metodologia de pontuação de crédito, conforme destacamos na seção 3.2.1 do presente trabalho. Além de consolidar-se como um órgão central na aplicação e interpretação da LGPD[523], considerado o direito à explicação, a ANPD também terá o desafio de realizar tal tarefa a partir de um amplo acúmulo jurisprudencial sobre a matéria, esclarecendo em que medida o judiciário reconheceu ou não a existência de um direito à explicação ou de obrigações de transparência amplas, e em que medida essa jurisprudência pode ou não ser recepcionada pela

---

[522] DONEDA, D. A Autoridade Nacional de Proteção de Dados e o Conselho Nacional de Proteção de Dados. *In*: MENDES, L. S.; DONEDA, D.; SARLET, I. W.; RODRIGUES JÚNIOR, O. L.; BIONI, B. (Coords.). *Tratado de Proteção de Dados Pessoais*. Rio de Janeiro: Forense, 2021. p. 465.

[523] A celebração do Acordo de Cooperação Técnica entre a Secretaria Nacional do Consumidor e a Autoridade Nacional de Proteção de Dados é demonstrativa do esforço da ANPD em firmar-se como órgão central de interpretação da LGPD já nos primeiros meses de sua atuação institucional. Nos termos do acordo celebrado, a Senacon ficará responsável por compartilhar com a ANPD informações coletadas sobre as reclamações de consumidores relacionadas à proteção de dados pessoais. A Autoridade, por sua vez, ficará responsável por fixar as interpretações necessárias à aplicação da LGPD nos casos concretos. Não há, todavia, ao menos até a data de publicação deste trabalho, nenhuma iniciativa de coordenação institucional articulada entre a ANPD e o Poder Judiciário. Cf. AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS. ANPD e Senacon assinam acordo de cooperação técnica. *ANPD*, 22 mar. 2021. Disponível em: https://www.gov.br/anpd/pt-br/assuntos/noticias/anpd-e-senacon-assinam-acordo-de-cooperacao-tecnica. Acesso em: 16 maio 2021.

LGPD.

Esse desafio de construção de uma interpretação sólida, sistemática, harmônica e uniforme em torno do direito à explicação pode ser ilustrado a partir de dois casos de grande repercussão envolvendo tensões entre demandas por maior transparência no contexto de sistemas algorítmicos e a defesa do segredo de negócio pelos agentes proprietários dos sistemas.

Em 2018, o Ministério Público do Rio de Janeiro moveu uma ação civil pública contra a empresa de comércio eletrônico Decolar, especializada na comercialização de passagens aéreas e pacotes de viagem. Na referida ação, o Ministério Público acusa a empresa "[...] pelas práticas de geo-blocking – bloqueio da oferta com base na origem geográfica do consumidor – e de geo-pricing – precificação diferenciada da oferta também com base na geolocalização."[524] Em síntese, o Ministério Público fundamentou seu pedido com base no entendimento de que a empresa

> [...] violou o direito brasileiro na medida em que se utilizou de tecnologia de informação para ativamente discriminar consumidores com base em sua origem geográfica e/ou nacionalidade para manipular as ofertas de hospedagem em hotéis, alterando o preço e a disponibilidade de ofertas conforme a origem do consumidor.[525]

O Superior Tribunal de Justiça reconheceu o interesse público em combater o preço discriminatório através de técnicas de geolocalização. Contudo, assegurou o segredo de justiça para proteger o algoritmo adotado pela empresa, para proteger a sua propriedade intelectual[526]. A proteção da propriedade intelectual fica bem clara na ementa na decisão:

> RECURSO ORDINÁRIO EM MANDADO DE SEGURANÇA. AÇÃO CIVIL PÚBLICA. DECRETAÇÃO DE SEGREDO DE JUSTIÇA. ILEGALIDADE. EXISTÊNCIA. GEODISCRIMINAÇÃO. GEO-PRICING. GEO-BLOCKING. PROCESSO COLETIVO. PUBLICIDADE. NECESSIDADE, COM RESGUARDO APENAS DOS DIREITOS DE PROPRIEDADE INTELECTUAL.
> 1. As práticas de "geodiscriminação" — discriminação geográfica de consumidores -, como o geo-pricing e o geo-blocking, desenvolvem-se no contexto da sociedade de risco e da informação, por intermédio de algoritmos computacionais, e — se comprovados — possuem a

[524] ZANATTA, R. A. F. Tutela coletiva e coletivização da proteção de dados pessoais. *In:* PALHARES, F. (Org.). *Temas Atuais de Proteção de Dados*. São Paulo: Thomson Reuters, 2020. p. 358.
[525] ZANATTA, R. A. F. Tutela coletiva e coletivização da proteção de dados pessoais. *In:* PALHARES, F. (Org.). *Temas Atuais de Proteção de Dados*. São Paulo: Thomson Reuters, 2020. p. 358.
[526] BRASIL. Superior Tribunal de Justiça. *Recurso em Mandado de Segurança nº 61.306 – RJ (2019/0199274-6)*. Relator: Min. Luís Felipe Salomão. Julgado em: 4/12/2019.

> potencialidade de causar danos a número incalculável de consumidores, em ofensa ao livre mercado e à ordem econômica.
> 2. O processo coletivo, instrumento vocacionado à tutela de situações deste jaez, é moldado pelo princípio da informação e publicidade adequadas (fair notice), segundo o qual a existência da ação coletiva deve ser comunicada aos membros do grupo.
> 3. A publicidade, erigida a norma fundamental pelo novo Código de Processo Civil (Art. 8º), garante transparência e torna efetivo o controle da atividade jurisdicional, motivo pelo qual também representa imperativo constitucional conforme se depreende do caput do art. 37 e do inciso IX do art. 93.
> 4. Não se desconhece que, em hipóteses excepcionais, é possível a decretação de sigilo de processos judiciais, conforme dispõe o art. 189 do CPC/2015. No entanto, na hipótese, tendo em vista os princípios que informam o processo coletivo e as garantias constitucionais e legais que socorrem os consumidores, o que na verdade atende o interesse público ou social é a publicidade do processo, que versa sobre possível prática de "geodiscriminação".
> *5. Outrossim, conforme requerido pelo próprio Ministério Público do Estado do Rio de Janeiro e com o escopo de, a um só tempo, resguardar o interesse público e preservar direitos de propriedade intelectual, considero razoável a manutenção do segredo de justiça tão somente no que diz respeito ao algoritmo adotado pela Decolar com Ltda. e à eventual perícia de informática relativa a tal algoritmo em toda a base de dados adotada para a operação do sistema de reservas eletrônicas.*
> 6. Recurso ordinário em mandado de segurança conhecido e parcialmente provido.[527] (grifos nossos).

Essa forte proteção ao segredo de negócio também se encontra presente no RE nº 13.047-36/RS, que tratou sobre obrigações de transparência na prática de pontuação de crédito, oportunamente analisado no Capítulo 3 do presente trabalho. No julgamento do recurso repetitivo, a corte reconheceu o interesse dos consumidores em terem acesso às justificativas para as decisões automatizadas às quais foram submetidos, mas estabelece que o interesse de agir face às instituições de crédito depende da recusa das mesmas em fornecer as motivações das avaliações de crédito que fizeram e que uma recusa de crédito seja fruto dessa avaliação[528]. Mais uma vez, a importância da proteção do segredo de negócio fica evidente na decisão da corte:

> APELAÇÃO CÍVEL. DIREITO PRIVADO NÃO ESPECIFICADO. AÇÃO CAUTELAR. SISTEMA DE PONTUAÇÃO. FALTA DE INTERESSE DE AGIR. NÃO CONFIGURADO. EXIBIÇÃO DE

[527] BRASIL. Superior Tribunal de Justiça. *Recurso em Mandado de Segurança nº 61.306 – RJ (2019/0199274-6)*. Relator: Min. Luís Felipe Salomão. Julgado em: 4/12/2019.
[528] BRASIL. Superior Tribunal de Justiça. *Recurso Especial nº 1304736 RS 2012/0031839-3*. Relator: Min. Luis Felipe Salomão. Data de Publicação: 30/03/2015. Disponível em: https://stj.jusbrasil.com.br/jurisprudencia/178798658/recurso-especial-resp-1304736-rs-2012-0031839-3. Acesso em: 17 ago. 2020.

> DOCUMENTOS. METODOLOGIA DE CÁLCULO. IMPOSSIBILIDADE. SEGREDO EMPRESARIAL. PRÉVIO E EXPRESSO CONSENTIMENTO. DESNECESSIDADE. PONTUAÇÃO DO CONSUMIDOR AVALIADO. DEVER DE EXIBIÇÃO. Existe interesse processual quando a parte tem necessidade de ir a juízo para alcançar a tutela pretendida. *De acordo com o entendimento emanado pelo Superior Tribunal de Justiça no julgamento do REsp 1.419.697-RS, a metodologia de cálculo da nota de risco de crédito constitui segredo empresarial, cujas fórmulas matemáticas e modelos estatísticos não precisam ser divulgados*. Ainda, não há necessidade de que haja prévio e expresso consentimento do consumidor avaliado. Com relação a pontuação, todavia, deve ser exibida, de acordo com o Código de Defesa do Consumidor e a Lei do Cadastro Positivo, contendo inclusive a indicação das fontes dos dados considerados na avaliação estatística, a fim de que o consumidor possa exercer um controle acerca da veracidade dos dados existentes sobre a sua pessoa, inclusive para poder ratificá-los ou melhorar a sua performance no mercado. APELO PROVIDO EM PARTE.[529] (grifo nosso).

No julgamento de uma recente ação trabalhista ajuizada em face da plataforma 99 Táxi, que opera no mercado de transporte individual, o entendimento do judiciário, desta vez representado pelo Tribunal Regional do Trabalho da 9ª Região, entendeu-se pela necessidade de dar maior proteção e peso à obrigação de transparência em detrimento da garantia do segredo de negócio. Em ação na justiça do trabalho em que a 99 Taxi figurava como ré (Ação Trabalhista nº 0000335-45.2020.5.09.0130)[530], discutia-se a existência de vínculo trabalhista entre uma motorista (reclamante) e a empresa (reclamada), sob os argumentos de que os motoristas se configuram como empregados conforme a definição do art. 3º da CLT[531]. O direito do trabalho avalia as condições efetivas das relações econômicas para além de disposições contratuais. Como boa parte das operações da empresa realiza-se de forma automatizada, incluindo boa parte do relacionamento com os motoristas, a discussão sobre a existência de relação de subordinação, elemento constitutivo de relações trabalhistas, dependia da análise dos algoritmos. Após a primeira audiência, o juiz então determinou uma auditoria no algoritmo da 99 Taxi, na qual deveriam ser analisadas (i) as condições em que as chamadas eram distribuídas pelo aplicativo, (ii) a forma em que se estabelecem os valores a serem repassados e (iii) a existência de condições

---

[529] RIO GRANDE DO SUL. Tribunal de Justiça do Estado do Rio Grande do Sul (6. Câmara Cível*). Apelação Cível nº 70059936971*. Data de Julgamento: 16/12/2014.

[530] BRASIL. Tribunal Regional do Trabalho da 9º Região. *Rito Sumaríssimo nº 0000335-45.2020.5.09.0130*. Relator: Leonardo Vieira Wandelli.

[531] BRASIL. Tribunal Regional do Trabalho da 9º Região. *Rito Sumaríssimo nº 0000335-45.2020.5.09.0130*. Relator: Leonardo Vieira Wandelli.

ou preferências no acesso e na distribuição dos chamados em função da avaliação do motorista ou do número de aceites de outras corridas.

Em respeito ao segredo de negócio, o juízo determinou que o processo corresse em segredo de justiça, e que apenas os resultados relativos aos pontos específicos sobre o relacionamento da empresa com os motoristas fosse juntada aos autos. O juízo negou a argumentação da defesa de que a revelação do algoritmo em auditoria representaria um dano às suas atividades pela exposição de seu segredo empresarial, visto que, sob sua ótica, a auditoria ocorreria em segredo de justiça, com seus resultados restritos ao juiz e à equipe técnica envolvida no caso:

> Outrossim, considerando-se que o processo passou a correr em segredo de justiça, conforme determinado na audiência em que proferida a decisão atacada, a ré não tem sucesso em indicar qualquer dano possível ao seu patrimônio imaterial (seja o segredo industrial, que estará protegido pelo sigilo judicial, seja o direito autoral do software, que somente seria violado em caso de apropriação por outrem mediante cópia utilizada em outro software), não se deduzindo qualquer dano da mera divulgação da ata de audiência contendo a decisão, que não desabona a ré nem expõe quaisquer dados privados ao público.[532]

De acordo com o órgão julgador, nenhuma outra prova, que não o código fonte do algoritmo, pode servir como elemento suficiente para sanar a questão da existência ou não de uma relação de subordinação característica de relações de emprego. Por determinação do tribunal, portanto, ficou determinada a realização de auditoria no algoritmo.[533] Ao final, contudo, em sede de conciliação, a empresa ofereceu uma proposta de acordo para que a autora da ação desistisse da demanda sem a necessidade de realização da auditoria. Não obstante o desfecho, esse caso aponta para um entendimento em direção oposta ao primeiro caso relatado acima,

---

[532] Ibidem, p. 16-17.

[533] Este também foi o entendimento do Tribunal Regional do Trabalho da 1ª Região em um caso envolvendo a Uber do Brasil Tecnologia LTDA., no qual se buscava analisar a existência de relação de emprego entre a empresa e um de seus motoristas, tendo sido determinada a realização de perícia no algoritmo. A perícia havia sido determinada pelo juízo de primeira instância, que entendeu que a prova era necessária à formação do livre convencimento do juiz. Alegando violação ao seu direito de propriedade intelectual e aos princípios da concorrência e da livre iniciativa, a empresa impetrou Mandado de Segurança requerendo a não realização da auditoria, argumentado que as provas necessárias à formação do convencimento do juiz poderiam ser supridas por vias alternativas e menos onerosas. Em síntese, a decisão do juízo foi a de determinar a realização de perícia no algoritmo, nos seguintes termos: "[...] apesar da proteção legal, acaso útil e necessária, a perícia técnica não pode ser inviabilizada no caso, tendo em vista que a própria legislação se encarrega de estabelecer parâmetros a compor proporcionalmente o direito coletivo à investigação e o direito individualizado à garantia do sigilo industrial." (BRASIL. Tribunal Regional do Trabalho da 1ª Região. *Mandado de Segurança Cível — Processo nº 0103519-41.2020.5.01.0000*. Julgado em: 29/04/2021).

envolvendo o MP-RJ e a Decolar.com, uma vez que o tribunal deu maior relevo às obrigações de transparência, determinando a abertura e realização de perícia técnica no algoritmo.

Não constitui nosso objetivo realizar uma análise exaustiva da jurisprudência envolvendo o direito à explicação e obrigações de transparência no contexto de sistemas automatizados, mas tão somente ilustrar, a partir de dois recortes muito específicos e pontuais, como a interpretação descentralizada do direito à explicação pelos tribunais pode impor desafios ao seu reconhecimento e à delimitação de sua exata extensão.

4.4 SÍNTESE DO CAPÍTULO

Embora optemos por debater cada um desses problemas relativos às limitações ao direito à explicação de forma separada, é importante ter em mente que eles estão essencialmente interrelacionados e possuem implicações mútuas quando pensamos na melhor estratégia de regulação. A divisão analítica é apenas uma etapa para possibilitar maior compreensão do problema. Os direitos à privacidade e a proteção de dados se inserem na intersecção entre o direito e o desenvolvimento de novas tecnologias e atividades econômicas. Mesmo naquilo que chamamos de capacidade técnica decorrente da tecnologia, o componente econômico ocupa um lugar central. Podemos em muitos casos afirmar que a utilização de sistemas não interpretáveis ou compreensíveis por seres humanos também faz parte de uma estratégia corporativa.[534] Os desafios para a implementação do direito à explicação, portanto, demandam esforços técnicos, teóricos, legislativos e jurisprudenciais. No próximo capítulo, iremos discorrer sobre como garantir o direito à explicação no contexto brasileiro, mesmo diante dessas limitações.

[534] RUDIN, C. Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead. *ArXiv [cs, stat]*, [*S. l.*], 2019. Disponível em: http://arxiv.org/abs/1811.10154. Acesso em: 10 ago. 2020.

5 A GARANTIA E A IMPLEMENTAÇÃO DO DIREITO À EXPLICAÇÃO NO ORDENAMENTO JURÍDICO BRASILEIRO NO CONTEXTO DE DECISÕES AUTOMATIZADAS

Já apresentamos as perspectivas teóricas sobre a existência do direito à explicação, discutimos porque ele é relevante no regime de proteção de dados, como este direito pode ser compreendido em diferentes ordenamentos jurídicos e os limites decorrentes do objeto que o direito visa regular, quais sejam, as tecnologias da informação ou do contexto social e jurídico. Uma vez que esperamos ter demonstrado que a mera obrigação de transparência não é suficiente para garantia da autodeterminação informacional nos contextos de procedimentos algorítmicos, podemos apresentar sobre a prática.

Neste capítulo discutiremos, em primeiro lugar, como garantir o direito à explicação a partir dos instrumentos jurídicos existentes na proteção de dados. O objetivo é apresentar como a "caixa de ferramentas" da LGPD pode ser utilizada nos contextos de decisões automatizadas. Vamos discutir cada um desses conceitos e sua implicação para o contexto das decisões automatizadas e do direito à explicação em específico.

Por fim, com base nos instrumentos jurídicos e técnicos para a implementação do direito à explicação, concluiremos o capítulo com orientações e considerações práticas para a implementação do direito a partir de uma abordagem centrada no sujeito de direitos e no devido processo informacional, a partir de cenários de aplicação.

5.1 IMPLEMENTAÇÃO A PARTIR DO DIREITO VIGENTE

Já discutimos neste trabalho como o direito à explicação aparece como um corolário do princípio da transparência e da autodeterminação informativa, e de como é importante compreendermos o tratamento de dados, e principalmente as decisões automatizadas, a partir de uma ótica do devido processo informacional. Essas conclusões baseiam-se em uma interpretação sistemática das regulações de proteção de dados estudadas. Essa interpretação é possível a partir da doutrina sobre a proteção de dados que vem se desenvolvendo nas últimas décadas e tem materializado os princípios e algumas normas em ordenamentos como o GDPR e a LGPD.

Uma vez que já apresentamos as questões doutrinárias mais relevantes, nesta seção discutiremos a implementação do direito à explicação no ordenamento brasileiro, utilizando subsidiariamente as experiências do direito europeu. As regulações supracitadas trazem diversas obrigações e conceitos, que servem como uma verdadeira caixa de ferramentas para a proteção de dados. O exercício de confrontar o processamento de dados com os conceitos e normas da LGPD é um processo que, em si mesmo, permite definir um quadro protetivo para os titulares de dados.

Nesta seção discutiremos os principais elementos da legislação de proteção de dados implicados no direito à explicação. Devemos destacar, contudo, que este trabalho não esgota a problemática da implementação de decisões automatizadas. Há questões importantes que podem ser levantadas quanto à definição de bases legais a partir do art. 7°, sobre transferências internacionais ou sobre o tratamento de dados sensíveis, dados de crianças e adolescentes, dentre outras. Por razões metodológicas não poderemos tratá-las neste trabalho, mas futuras investigações sobre o tema poderão contribuir sobremaneira para o desenvolvimento da matéria.

Em primeiro lugar abordaremos o regime de responsabilização da LGPD, discutindo como o direito à explicação deve ser compreendido nas relações entre os agentes de tratamento e as principais implicações para a implementação de decisões automatizadas nos fluxos de dados complexos que contemplam diversos agentes.

Em seguida, discutiremos o direito à explicação em relação às demais obrigações de transparência previstas na LGPD, diferenciando as obrigações de transparência ativas ou passivas. Então, apresentamos a problemática do regime da a*ccountability* estabelecido pela lei nas aplicações de decisões automatizadas e do direito à explicação, discutindo brevemente o princípio e argumentando sobre quais são os atores aos quais as organizações devem prestar contas.

Reservamos uma seção para as discussões sobre os relatórios de impacto e as diferentes propostas de avaliações que emergiram no campo de estudos sobre algoritmos e inteligência artificial. E, além disso, a partir do conceito de *privacy by design* adotado pela LGPD, discutiremos como conceito de *explainability by design* pode contribuir com o desenvolvimento dessas tecnologias. Por fim, reservamos uma seção sobre as auditorias, na qual discutiremos as prerrogativas das autoridades nacionais para realização de auditorias e sobre a possibilidade de uso desse tipo de ferramentas pelas próprias organizações como forma de desenvolvimento e

monitoramento de suas aplicações.

5.1.1 Os Agentes de tratamento

Um dos elementos mais relevantes nas regulações de dados contemporâneas como o GDPR e a LGPD diz respeito à delimitação de responsabilidades pelo tratamento de dados. Nesse sentido, na diretiva 95/46 surge o conceito de controlador (*controller*). Essa figura jurídica aparece como elemento central na organização da proteção de dados. Naquele contexto, o conceito servia para diferenciar, de um lado, as organizações que deveriam ser as responsáveis pelos danos oriundos do tratamento de dados, e de outro, aquelas organizações que realizam o tratamento a pedido, ou em nome, de outrem, os quais seriam denominados operadores (*processors*).

A diretiva impunha então uma série de obrigações aos controladores de dados pessoais. E, como consequência, criava um regime de responsabilização diferencial em que a responsabilidade centrava-se no agente ao qual cabiam as decisões relativas ao processamento de dados, visto que a diretiva ainda não estipulava responsabilidades específicas para os operadores[535]. Numa consulta realizada pelo Working Party 29 em 2010 para avaliação da diretiva, algumas organizações argumentaram que a diferenciação não correspondia à complexidade da cadeia organizacional do fluxo de tratamentos de dados pessoais. Sugeriu-se que a divisão de responsabilidades deveria ser mais flexível, de acordo com a responsabilidade específica de cada parte no processamento[536].

Apesar das críticas e sugestões pelo fim da diferenciação, o GDPR manteve os dois conceitos e estabeleceu critérios mais claros para a responsabilização pelo processamento de dados, como a previsão da categoria de “co-controladores”, a criação da figura do *Data Protection Officer* (DPO), dentro outras. Os conceitos também foram adotados pela LGPD, classificados como “agentes de tratamento”, categoria que dá nome ao capítulo VI da lei.

Há diferenças significativas entre o regime da LGPD e do GDPR em relação ao

---

[535] VAN ALSENOY, B. Liability under EU Data Protection Law: From Directive 95/46 to the General Data Protection Regulation. *JIPITEC*, [*S. l.*], v. 7, n. 3, 2017. Disponível em: http://www.jipitec.eu/issues/jipitec-7-3-2016/4506. Acesso em: 21 jun; 2021.

[536] VAN ALSENOY, B. Liability under EU Data Protection Law: From Directive 95/46 to the General Data Protection Regulation. *JIPITEC*, [*S. l.*], v. 7, n. 3, 2017. Disponível em: http://www.jipitec.eu/issues/jipitec-7-3-2016/4506. Acesso em: 21 jun; 2021.

regime específico de responsabilidade civil[537]. No entanto, as duas regulações possuem em comum a preocupação de que os danos aos titulares de dados possam encontrar o agente para responsabilização de uma forma em que a comprovação de culpa seja menos relevante, sem, contudo, estabelecer-se um regime estrito de responsabilidade objetiva[538].

O Art. 5º da LGPD define o controlador como “pessoa natural ou jurídica, de direito público ou privado, a quem compete as decisões referentes ao tratamento de dados pessoais”. O operador é definido, nos mesmos termos, como aquele que realiza o tratamento de dados em nome do controlador. Nesse sentido, o primeiro é aquele a quem cabe decidir, por competência legal pré-definida ou da capacidade determinante de determinada cadeia de processamento de dados.

É importante destacar que apesar da complexidade do processamento de dados, a legislação realizou um recorte. A regulação fez uma escolha de responsabilização naquele agente mais influente na cadeia. O controlador aparece então como um dos elementos principais da regulação: a palavra aparece mais de 50 vezes na LGPD. O operador aparece apenas 14 vezes no texto. A responsabilidade do operador é mais restrita: é solidário nos casos em que descumprir a legislação de proteção de dados ou agir de forma diversa do estipulado pelo controlador.

Assim como no contexto Europeu, houve mobilização do setor privado contra o regime da responsabilidade objetiva do controlador e contra a responsabilidade solidária. Defendeu-se o regime de responsabilidades separadas e flexíveis de acordo com cada tratamento de dados. Embora não tenha havido um regime estrito de responsabilidade objetiva, o regime da LGPD é menos flexível do que o defendido pelos agentes privados.[539]

Rafael Zanatta demonstra como os agentes de tratamento na LGPD representam uma importação parcial da legislação europeia. A lei brasileira estipula que nas hipóteses de violação dos direitos do consumidor, deverá ser utilizado o regime do direito consumerista. Neste sentido, observamos um regime misto de responsabilidade que resulta da integração da LGPD e do Código de Defesa do

---

[537] ZANATTA, R. A. F. Agentes De Tratamento De Dados, Atribuições e Diálogo Com O Código De Defesa Do Consumidor. *Revista dos Tribunais*, [*S. l.*], n. 1009, supl. Caderno Especial, nov. 2019. p. 188.

[538] ZANATTA, R. A. F. Agentes De Tratamento De Dados, Atribuições e Diálogo Com O Código De Defesa Do Consumidor. *Revista dos Tribunais*, [*S. l.*], n. 1009, supl. Caderno Especial, nov. 2019.

[539] ZANATTA, R. A. F. Agentes De Tratamento De Dados, Atribuições e Diálogo Com O Código De Defesa Do Consumidor. *Revista dos Tribunais*, [*S. l.*], n. 1009, supl. Caderno Especial, nov. 2019.

Consumidor. E nesse sentido, em alguns casos, poderemos observar um regime de responsabilidade objetiva.

A divisão entre operadores e controladores implica em questões relevantes de aplicação, visto que o conceito rígido não corresponde à fluidez das situações práticas. Essa afirmação é ainda mais verdadeira no contexto das decisões automatizadas. Se pensarmos no setor tecnológico, nos casos em que uma organização decide automatizar alguns de seus processos, é comum que terceiros forneçam ou desenvolvam tais aplicações. Algumas delas operam sob licença, desenvolvidas a partir de fontes próprias do terceiro. Outras vezes são desenvolvidas sob demanda, por encomenda, a partir de dados da organização contratante. Em ambos os casos pode ser difícil definir na prática qual seria o agente responsável por cada uma das decisões sobre o tratamento de dados.

No caso em que uma organização contrata o desenvolvimento de *software*, ainda que seja responsável pela definição dos dados coletados e das bases legais para o tratamento, há uma série de aspectos técnicos relevantes que poderão ocorrer a cargo dos operadores de dados. As metodologias de desenvolvimento do algoritmo, as métricas utilizadas para testar a eficiência, eficácia e porcentagens de erro da aplicação etc. Esses elementos possuem implicações numa decisão automática e, em alguns casos, significam informações importantes para os titulares de dados.

Essa condição impõe algumas consequências relevantes. Em primeiro lugar é importante notar como tal condição impõe alguns cuidados para definição das posições e responsabilidades nas relações contratuais. O dever de cuidado do operador na escolha de seus parceiros comerciais e operacionais demanda muita diligência no processo de avaliação sobre a capacidade do operador de cumprir a legislação.

O GDPR prevê expressamente a existência de relações de *joint controllers* ou co-controladores. Essa relação deve estabelecer de forma transparente as respectivas responsabilidades de cada parte, de forma que cada controlador tenha as bases legais, o dever de cumprir as obrigações e de garantir os direitos dos titulares. Neste sentido, as Guidelines do EDPB recomendam que sejam utilizadas formas contratuais[540].

[540] EUROPEAN DATA PROTECTION BOARD. *Guidelines 07/2020 on the concepts of controller and processor in the GDPR*. [*S. l.*]: EDPB, 2020. Disponível em:

Embora não prevista na legislação brasileira, a situação pode ocorrer com frequência. E essa divisão é muito relevante para a discussão sobre o direito à explicação. No contexto da divisão de responsabilidades, é preciso definir, em disposições contratuais, aquele a quem cabe a tarefa de explicar. Pelo texto do § 1° do art. 20 inicialmente podemos afirmar que tal obrigação cabe ao controlador[541]. No entanto, isto não impede que uma requisição de informações seja requerida a um operador, pode-se questionar se tal obrigação também não deve se aplicar também a ele.

No complexo fluxo de dados, onde o contato com os titulares de dados pode ocorrer de várias formas e por diferentes agentes e as escolhas sobre os processamentos de dados são fluídas, a discussão sobre quem é responsável pela explicação ganha contornos a depender da casuística. Mas, a despeito das especificidades de cada aplicação, é certo que o cuidado e os esforços de adequação à proteção de dados deverão envolver uma visão global do processamento de dados, o que inclui preocupar-se com os processamentos realizados internamente ou por terceiros.

Neste sentido, tão ou mais importante do que a mera definição de operadores e controladores é a avaliação cuidadosa das medidas de segurança, do fluxo de dados e dos princípios da lei para que os direitos dos titulares sejam garantidos em toda a cadeia. É preciso considerar os deveres oriundos do princípio da transparência e os direitos previstos na legislação.

A transparência aparece como um princípio no art. 6º, VI, da LGPD. Nesse sentido, o fornecimento de informações claras, precisas e facilmente acessíveis é um dos elementos centrais da proteção de dados. Isto posto, podemos afirmar que as obrigações de transparência cabem a qualquer agente de tratamento envolvido no processamento de dados. Além disso, o direito de acesso disciplinado pelo art. 9º não diferencia entre operadores e controladores. A menção ao controlador ocorre para obrigar a sua identificação, bem como uma forma de contato, o que pressupõe que tal

---

https://edpb.europa.eu/sites/default/files/consultation/edpb_guidelines_202007_controllerprocessor_en.pdf. Acesso em: 3 maio 2021.

[541] "Art. 20. [...] § 1º O controlador deverá fornecer, sempre que solicitadas, informações claras e adequadas a respeito dos critérios e dos procedimentos utilizados para a decisão automatizada, observados os segredos comercial e industrial." (BRASIL. *Lei nº 13.709, de 14 de agosto de 2018*. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF: Presidência da República, 14 ago. 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015- 2018/2018/lei/L13709.htm. Acesso em: 07. nov. 2020).

obrigação pode ser realizada por terceiros.

A previsão de que tal obrigação pode ser realizada por terceiros, no caso, operadores, não torna as obrigações do controlador completamente delegáveis. A responsabilidade pelas escolhas e decisões do tratamento continuam sob a esfera do controlador de forma que o operador se responsabilizaria apenas pelas ações em desconformidade com a legislação ou as orientações do controlador. Nesse sentido, o próprio art. 9 ainda recomenda que sejam fornecidas informações acerca do uso compartilhado e da responsabilização dos agentes que realizaram o tratamento.

O princípio da transparência impõe que as informações sejam úteis e que contenham os conteúdos mais relevantes, ou que afetem a esfera de direitos do titular. Nesse sentido, o art. 9º obriga que se forneçam informações sobre todos os direitos do titular com menção explícita aos direitos previstos no art. 18. Embora o direito à explicação não esteja previsto no Art.18, caso haja decisões automatizadas, isto deve incluir os direitos de revisão e de explicação.

A interpretação mais razoável da leitura sistemática destes artigos parece apontar para uma leitura de que o fornecimento de informações e o cumprimento das obrigações de transparência pode ser realizado por controladores e operadores. No entanto, teremos um regime de responsabilização dos controladores, respondendo o operador solidariamente em caso de descumprimento contratual ou ato ilícito. No entanto, nos casos em que haja expectativas razoáveis de que um agente de tratamento forneça explicações, nos casos em que tenha autonomia e controle sobre aspectos relevantes do processamento, mesmo que exerça o papel de operador, pode-se concluir pela sua responsabilização, caso se entenda que tenha agido em desacordo com o regramento da proteção de dados.

Para as decisões automatizadas, conforme o § 1º do art. 20, contudo, encontramos uma menção direta ao controlador. Segundo o dispositivo, cabe ao controlador fornecer informações claras e adequadas sobre os critérios e os procedimentos utilizados para a decisão automatizada, observados os segredos comercial e industrial. Os segredos comerciais e industriais novamente adicionam uma camada de complexidade na questão.

Os segredos comerciais e industriais são perfeitamente oponíveis nas relações empresariais. Nos casos em que algoritmos de terceiros são utilizados para operações de decisões automatizadas, contudo, alguns impasses podem emergir. Nesse sentido, para que o agente de tratamento possa fornecer informações claras e adequadas

sobre o processamento e os critérios utilizados no processamento, é preciso que o parceiro contratual forneça essas informações relevantes, que devem incluir especificidades técnicas do sistema. A depender do nível de informações necessárias para esclarecimento do titular de dados, esse processo pode pôr em risco aspectos da estratégia comercial das organizações.

É bem provável que algoritmos tenham utilizado, em seu desenvolvimento, dados de terceiros, mas que, na relação específica de uma decisão automatizada, em relação aos dados de um titular, a empresa fornecedora do serviço seja uma operadora de dados pessoais. Nesses casos, é razoável concluir que operadores de dados devem ser capazes de fornecer elementos suficientes para a explicação dos critérios envolvidos e dos principais processos envolvidos nas decisões automatizadas.

É de se imaginar que tal necessidade permeia as negociações comerciais dos setores tecnológicos, e que contratantes que figurem como controladoras de dados pessoais em operações de decisões automatizadas tenham como precaução a exigência de explicabilidade ante a contratação de sistemas de terceiros. A necessidade de transparência demanda uma postura ativa dos desenvolvedores de aplicações, que devem demonstrar que suas aplicações não possuem viés social, que não são discriminatórios, que são eficientes, precisos e seguros. Tais características, como demonstrado pela discussão dessa tese, integram o corpo regulatório da proteção de dados pessoais.

Neste sentido, é importante que todas as organizações cujas atividades principais demandam processamento de dados para decisões automatizadas, como perfilização, *scoring*, mediação de relações comerciais, atendimentos etc. devem estar atentas para a proteção de dados como um todo. Uma visão que encare posições de controlador e operador como realidades estanques pode resultar em riscos jurídicos consideráveis. Neste sentido, as discussões deste capítulo devem ser compreendidas como reflexões direcionadas a qualquer agente de tratamento de dados pessoais.

5.1.2 Os direitos morais dos titulares de dados e obrigações de transparência: acesso, explicação e revisão

A LGPD e outras leis setoriais, como o Marco Civil da Internet, o Código de

Defesa do Consumidor e a Lei do Cadastro Positivo, garantem aos titulares dos dados o direito de acesso à integralidade dos dados pessoais que uma organização, pública e privada, detém sobre ele. Todavia, o mero fornecimento desses dados, como já explicado anteriormente, provavelmente não permitirá ao titular ter um conhecimento efetivo sobre a forma e para o que eles são tratados, muito menos compreender os seus possíveis impactos na sua vida. Por isso que esse e outros direitos são acompanhados de obrigações de transparência ativa e passiva, que estabelecem o dever de fornecer ao titular um mínimo de informações sobre tratamento dos seus dados, incluindo os fins para os quais estes serão utilizados.

A palavra *transparência* é utilizada amplamente em vários contextos. Já discutimos como a mera obrigação de informar é inefetiva quando desacompanhada de outros mecanismos de controle. Contudo, é impossível negar o seu papel central na regulação de privacidade e proteção de dados. É a partir da obrigação de transparência, como elemento de um devido processo informacional, que se pode argumentar em favor de um direito à explicação.

Resgatar e entender o conceito de transparência, portanto, é importante, pois, em geral, a transparência é usada como uma forma de descrever um fenômeno, mas raramente é abordada como algo a ser explicado. Como já abordado no item 2.2.2, a possibilidade de ver um sistema não significa necessariamente entender seu funcionamento e governá-lo. Portanto, a instrumentalização da transparência não necessariamente significa discernimento da tecnologia empregada. Nesse sentido, cabe apresentarmos brevemente o histórico do conceito e como devemos interpretá-lo no atual corpo regulatório da proteção de dados.

O conceito de transparência só tomou a forma atual no contexto do estado democrático de direito na década de 1990. No artigo “Whats is transparency”[542], Ball analisa a evolução da definição de transparência na língua inglesa. A autora rastreia o significado de transparência desde seu uso por organizações não governamentais e supranacionais até seu uso na literatura de relações internacionais, organizações sem fins lucrativos, políticas públicas e administração. Segundo a autora, a presença da palavra transparência em documentos organizacionais inicia-se na década de 1990

[542] BALL, C. What Is Transparency? *Public Integrity*, [*S. l.*], v. 11, n. 4, p. 293–308, 2009. Disponível em: https://doi.org/10.2753/PIN1099-9922110400. Acesso em: 21 jun. 2021.

em torno da criação da União Europeia[543]. Mas provavelmente a escolha do nome "Transparência Internacional" para uma ONG e as atividades subsequentes dessa organização ajudaram a definir a palavra para o público e o mundo acadêmico.

Sem surpresa, com a internet e a digitalização do mercado financeiro, o termo "transparência" ganhou cada vez mais destaque. Na década de 1990, com a transferência financeira em tempo real, tornou-se possível altas movimentações de forma rápida de um país para outro, criando cenários propícios para corridas de moeda, contribuindo para crises financeiras. A governança fiscal e monetária opaca, com Bancos Centrais e Tesouros operando com falta de transparência, levavam gestores a confundir os dados ou simplesmente esconderem informações até que fosse tarde demais e o desastre fosse inevitável[544].

Como resposta, organismos multilaterais adotaram a existência de mecanismos de transparência como critério de empréstimos a organizações e países. Um dos resultados foi a explosão das leis de liberdade e acesso à informação[545]. Como as organizações multilaterais obrigaram os governos a adotar mecanismos de transparência, os cidadãos e os políticos também buscaram maior transparência de seus parceiros. Aos poucos as obrigações de transparência se tornaram mais multidirecionais.

O conceito, de origem política, transcende o debate travado na esfera pública e começa a ser utilizado também nas relações privadas. No contexto da disciplina da proteção de dados assume um papel central, visto que na falta de transparência, não se pode avaliar nem a qualidade dos dados nem a qualidade do processamento. Daí o diagnóstico preocupante de que os resultados algorítmicos, na atualidade, correspondem a uma verdadeira *black box*[546]*,* a uma quase total ausência de transparência*.*

O princípio da transparência, no âmbito da proteção de dados, surge como uma forma de empoderar o titular dos dados pessoais, possibilitando que possa ter

---

[543] LODGE, J. Transparency and Democratic Legitimacy. *JCMS: Journal of Commom Market Studie***s**, [*S. l.*], 1994. Disponível em: https://doi.org/10.1111/j.1468-5965.1994.tb00501.x. Acesso em: 4 dez. 2020.

[544] MICHENER, G.; BERSCH, K. Identifying Transparency. *Inf. Polity*, v. 18, n. 3, p. 233-242, 2013. Disponível em: https://doi.org/10.3233/IP-130299. Acesso em: 21 jun.2021.

[545] ACKERMAN, J. M.; SANDOVAL-BALLESTEROS, I. E. The Global Explosion of Freedom of Information Laws. *Administrative Law Review*, [*S. l.*], v. 58, n. 1, p. 85–130, 2006.

[546] PASQUALE, F. *The Black box society:* the secret algorithms that control money and information. First Harvard University Press paperback edition ed. Cambridge, Massachusetts; London, England: Harvard University Press, 2015.

informações claras, precisas e facilmente acessíveis sobre as atividades de tratamento de dados que está sujeito[547]. Na União Europeia, desde a diretiva e dos trabalhos do WP29, a transparência já aparecia com um quadro mais bem definido para sua implementação considerando elementos necessários, as informações a serem fornecidas e forma de divulgação [548].

Para fins conceituais, de forma a compreendermos a instrumentalização da transparência, podemos dividi-la em duas dimensões principais: a visibilidade e inferibilidade[549]. A primeira dimensão representa o grau em que a informação é completa e facilmente localizada (visível) e até que ponto ela pode ser usada para tirar conclusões precisas (inferível). Por exemplo, embora as fórmulas matemáticas que descrevem a inflação possam ser visíveis, nem todas as pessoas compreenderão matemática não simplificada; portanto, a inferibilidade depende do público-alvo.

Na LGPD a transparência aparece como princípio e é definida como "garantia, aos titulares, de informações claras, precisas e facilmente acessíveis sobre a realização do tratamento e os respectivos agentes de tratamento, observados os segredos comercial e industrial" (art. 6º, VI, da LGPD). Portanto, busca garantir que as informações ou comunicações relacionadas com o tratamento dos dados pessoais sejam de fácil acesso e compreensão, e formuladas numa linguagem clara e simples.

É importante destacar que mesmo que a informação sobre o tratamento de dados apresente relativo grau de compreensibilidade, ela não é completa se não for de fácil acesso aos titulares. Nesse sentido, a facilidade de acesso às informações sobre as atividades de tratamento também é um requisito fundamental para a caracterização da transparência no âmbito da proteção de dados pessoais.

A Lei prevê um conjunto de ferramentas que se traduzem em mecanismos que aprofundam obrigações de transparência. A transparência deve ocorrer, de acordo com as regulações de proteção de dados, de forma ativa, pela obrigação da difusão sistematizada de informações sobre o tratamento de dados e através de uma política de proteção de dados clara. Podemos incluir nesse rol, os mecanismos que facilitem o titular entender as etapas de processamento dos seus dados, disponibilizadas de

---

[547] BENNET, C. Regulating privacy: data protection and public policy in Europe and the United States. Ithaca, New York: Cornell University Press, 1992. p. 153-192.

[548] UNIÃO EUROPEIA. *WP 260rev.01* — Guidelines on Transparency under Regulation 2016/679. [*S. l.*]: WP29, 2018. Disponível em: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=622227. Acesso em; 21 jun. 2021.

[549] MICHENER, G.; BERSCH, K. Identifying Transparency. *Inf. Polity*, v. 18, n. 3, p. 233-242, 2013. Disponível em: https://doi.org/10.3233/IP-130299. Acesso em: 21 jun.2021.

forma clara, adequada e ostensiva.

Por outro lado, há uma outra categoria de práticas de transparência que podemos chamar de passivas — de forma que se pode deduzir outro princípio, o do livre acesso —, que se refere à obrigação do controlador em conceder, a todos os titulares que o requeiram, o acesso tempestivo aos seus dados pessoais na íntegra, salvo aqueles que estiverem legalmente protegidos por motivo de segredos comercial e industrial, ou por limitações técnicas que inviabilizem o seu fornecimento.

Dessa forma, uma diferença central entre a transparência e o acesso é que a primeira só pode descrever as informações sobre coleta de dados em termos mais abstratos. Portanto, a transparência a *priori* descreve as categorias de dados, a forma de processamento, além de trazer outras informações ao titular. O direito de acesso, transparência a *posteriori,* pode ser utilizado como forma de conferir se a transparência *a priori* está correta, já que o titular possui acesso a integralidade dos dados. Por exemplo, seguindo o princípio da transparência, uma empresa pode dizer que coleta nomes e, com a solicitação apresentada com base no direito de acesso, podemos saber, por exemplo, que o nome coletado era João da Silva.

Dessa forma, o acesso pode ser uma ferramenta central para o empoderamento do titular dos dados, sendo importante para diminuir a assimetria de poder informacional em uma sociedade cada vez mais moldada pelo uso de dados. Porém, essa arquitetura do direito/princípio de acesso só funciona quando este se encontra apoiado numa política de transparência ampla, além de precisar ser apoiada por altos níveis de *accountability*, *compliance* e fiscalização, que discutimos no item 2.1.2 deste trabalho, sobre os limites da transparência[550].

Portanto, o direito de acesso possui um papel fundamental na legislação brasileira. Apenas o direito de acesso permite ao titular dos dados exercer outros direitos, como a retificação, apagamento e, de certa forma, também o direito à revisão de decisões automatizadas e a efetividade de uma explicação. Mas, por outro lado, o acesso por si só não serve como forma suficiente para que o titular dos dados possua controle sobre fluxos e procedimentos que afetam a sua personalidade, pois o mero fornecimento desses dados provavelmente não permitirá ao titular ter um conhecimento efetivo sobre a forma, como e para o que eles são tratados.

Resgatando a comparação da transparência pública, somente cidadãos

[550] ANANNY, M.; CRAWFORD, K. Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, v. 20, n. 3, p. 973–989, 2016.

informados podem fazer julgamentos políticos informados com relação a um governo que, em uma sociedade democrática, deveria estar sob seu controle. A justificativa para ter o direito de acesso é muito semelhante. Porém, as desigualdades de informação e poder, vividas pelos cidadãos, impedem o efetivo controle da administração pública. É possível afirmar que o mesmo problema pode ocorrer com os direitos de acesso no contexto do processamento de dados pessoais[551].

Além disso, embora o direito de acesso tenha sido codificado de forma que deva ser relativamente fácil para o cidadão executar, por ter um baixo nível de requisitos formais para a solicitação, obter uma imagem clara das práticas de dados por meio do exercício do direito de acesso ainda é muito difícil[552]. As organizações precisam limitar a acessibilidade à informação de muitas maneiras diferentes por conta de estratégias corporativas ou limitações técnicas, conforme discutimos no item 4.1 deste trabalho.

Por fim, mesmo quando as obrigações de transparência passiva e ativa garantam aos titulares dos dados o direito de acesso a integralidade dos dados pessoais que uma organização, pública e privada, detém sobre ele, o mero fornecimento desses dados, como já discorrido em extensão, provavelmente não permitirá ao titular ter um conhecimento efetivo como seus dados são tratados, muito menos em sistemas complexos automatizados.

Isso porque, essas situações encontram limitações na cognição do próprio titular para compreender as informações que lhe são repassadas, principalmente as relativas à lógica subjacente ao processamento dos dados, conforme discutimos no item 4.2 deste trabalho. Portanto, o direito de acesso, em si, mesmo atrelado às obrigações de transparência, pode não ser suficiente para garantir o efetivo controle sobre o fluxo de dados e a autodeterminação informativa, de forma que se desdobre em um verdadeiro direito à explicação.

5.1.3 *Accountability* e responsabilidade demonstrável

O princípio da *accountability* é uma constante na maioria das regulações de

[551] MAHIEU, R.; AUSLOOS, J. *Recognising and Enabling the Collective Dimension of the GDPR and the Right of Access*. *LawArXiv*, 2 jul. 2020. Disponível em: https://doi.org/10.31228/osf.io/b5dwm. Acesso em: 9 dez. 2020.

[552] MAHIEU, R.; ASGHARI, H.; VAN EETEN, M. *Collectively Exercising the Right of Access*: Individual Effort, Societal Effect. Rochester, NY: Social Science Research Network, 2017. SSRN Scholarly Paper. Disponível em: https://doi.org/10.2139/ssrn.3107292. Acesso em: 9 dez. 2020.

proteção de dados. De forma apressada e com as devidas ressalvas, é possível estabelecer um caminho de desenvolvimento do princípio a começar pelas *Guidelines* da OCDE para proteção de dados, estabelecidas em 1980, passando pelo *Personal Information Protection and Electronic Documents Act* (PIPEDA) em 2000 no Canadá- que elenca *accountability* entre seus princípios —, pelo *Asia-Pacific Economic Cooperation* (APEC) Privacy Framework de 2005 e pelo *Accountability Project* do *Centre for Information Policy Leadership* (CIPL) concluído em 2009 e que contou com várias rodadas de discussão[553], culminando com a Diretiva 95/46, o GDPR e consequentemente a LGPD.

Tais recomendações e documentos emergem em um contexto de mudança da abordagem regulatória nas mais diferentes áreas. A chamada *nova administração pública* teve forte influência em algumas áreas emergentes do direito como o direito ambiental e a proteção de dados. Essa abordagem regulatória privilegia mecanismos de autorregulação e de ações tomadas pelos próprios agentes econômicos em suas atividades. Tal proposta permite um caminho para o cumprimento da legislação ao mesmo tempo em que se preserva a livre iniciativa e a liberdade de empreender[554].

O termo é amplamente discutido e assume vários significados diferentes a depender do contexto ou do domínio de aplicação. De maneira simplificada, contudo, podemos afirmar que o conceito de *accountability* remete a necessidade de prestação de contas sobre determinada atividade. Nesse sentido, pressupõe ao menos dois sujeitos: um sujeito que presta contas e outro sujeito a quem essa prestação é endereçada. Essa obrigação ocorre porque há uma relação de poder entre esses dois sujeitos. O conceito é utilizado para prever mecanismos de controle das atividades daquele que exerce o poder, seja do setor público sobre os cidadãos, seja de entes privados que exerçam posições de poder numa relação privada onde há alguma modalidade de assimetria negocial, econômica ou informacional.

O conceito consolidou-se nas legislações da chamada 4ª geração da proteção de dados pessoais, que vão além das antigas abordagens de proteção de dados, que privilegiavam a escolha e controle do fluxo de dados pelo próprio sujeito, pois essas,

---

[553] ALHADEFF, J.; VAN ALSENOY, B.; DUMORTIER, J. The Accountability Principle in Data Protection Regulation: Origin, Development and Future Directions. *In*: GUAGNIN, D. *et al.* (org.). *Managing Privacy through Accountability*. London: Palgrave Macmillan UK, 2012. p. 49–82. Disponível em: https://doi.org/10.1057/9781137032225_4. Acesso em: 16 dez. 2020.

[554] MULGAN, R. Issues of Accountability. *In*: MULGAN, R. (Org.). *Holding Power to Account*: Accountability in Modern Democracies. London: Palgrave Macmillan UK, 2003. p. 1–35. Disponível em: https://doi.org/10.1057/9781403943835_1. Acesso em: 13 nov. 2020.

em muitos cenários, mostravam-se insuficientes[555]. O desenvolvimento das tecnologias de informação e comunicação consolidava uma *commoditização* dos dados pessoais a partir de uma rede complexa de agentes de tratamento que extrapola a capacidade de agência de indivíduos isoladamente[556]. A escala e a intensidade do fluxo de dados pessoais tornaram a ideia de um pleno controle dos dados pessoais pelo titular inviável[557].

A consequência esperada pelos reguladores com tal abordagem foi fazer com que a responsabilidade pela proteção de dados se deslocasse do sujeito (titular dos dados) para os agentes do tratamento de dados. Tais operações de tratamento passam a operar mediante um regime de confiança, conferida ao agente de tratamento, através do estabelecimento de obrigações e práticas que possam garantir tal confiança. O conceito de *accountability*, então, cumpre o papel de incentivar o estabelecimento, pelos agentes de tratamento, de tais práticas para a proteção de dados.

O conceito de risco também se torna um elemento central para este regime regulatório, visto que é a partir dos riscos encontrados, que se avaliam as medidas necessárias, a partir de um cálculo de proporcionalidade, para os mitigar. A avaliação desse risco, em geral, é de responsabilidade dos próprios agentes de tratamento, que responderam pela qualidade dessa avaliação e pela efetividade das medidas tomadas. A LGPD estabelece alguns instrumentos específicos para essas avaliações, como os relatórios de impacto, dos quais trataremos a seguir. Mas, além dessas previsões específicas, o princípio da *accountability* estabelece que os agentes de tratamento são os responsáveis não apenas por tomar medidas eficientes para a proteção de dados pessoais, mas por ser capaz de demonstrá-las[558].

Embora o conceito tenha sido adotado no âmbito na proteção de dados desde, pelo menos, os anos 90, uma série de desafios fizeram-se presentes para a sua efetividade. Há uma questão relativa à capacidade dos órgãos reguladores e

---

[555] DONEDA, D. *Da privacidade à proteção de dados pessoais:* fundamentos da Lei Geral de Proteção de dados. 2. ed. São Paulo: Thomson Reuters Brasil, 2019. p. 172–180.
[556] BIONI, B. R. *Proteção de dados pessoais:* a função e os limites do consentimento. 1. ed. Rio de Janeiro: Forense, 2019. p. 112.
[557] ALHADEFF, J.; VAN ALSENOY, B.; DUMORTIER, J. The Accountability Principle in Data Protection Regulation: Origin, Development and Future Directions. *In*: GUAGNIN, D. *et al.* (org.). *Managing Privacy through Accountability*. London: Palgrave Macmillan UK, 2012. p. 49–82. Disponível em: https://doi.org/10.1057/9781137032225_4. Acesso em: 16 dez. 2020.
[558] GELLERT, R. Understanding data protection as risk regulation. *Journal of Internet Law*, [*S. l.*], p. 3–15, 2015.

fiscalizadores acompanharem o desenvolvimento rápido da tecnologia. Nesse sentido, a previsão e o uso do conceito, em si, não garantem a sua efetividade, visto que a autorregulação apresenta limitações importantes [559]. É preciso que o estado tenha poder de *enforcement* para que os agentes econômicos implementem medidas de fato.

Em 2010, um documento do Article 29 Working Party[560]**,** dedicado ao problema da *accountability*, apontava o problema da dificuldade de implementar os princípios na prática. O princípio precisa ser acompanhado de mecanismos legais de responsabilização de forma que o estado precisa de uma estrutura de fiscalização e de poder sancionatório.

Nesse sentido, o GDPR foi visto como uma evolução importante, ao presentear as agências reguladoras com poder sancionatório muito maior. Anteriormente conhecidas como "cães de guarda sem dentes", as autoridades nacionais de proteção de dados não possuíam mecanismos efetivos apesar de regras e princípios protetivos. A previsão de penalidades do regulamento trouxe um incentivo significativo para a implementação das práticas de *accountability* e prestação de contas[561].

Na LGPD, o princípio aparece no inciso X do art. 6°, como "responsabilização e prestação de contas". O artigo, que trata dos princípios para o tratamento de dados, foi originado no anteprojeto de Autoria do Ministério da Justiça, posteriormente apensado ao PL nº 4060 de 2012, que tramitava na Câmara[562]. Enquanto se discutia esse tema no projeto de lei do poder executivo, não havia uma previsão da prestação de contas e da responsabilização como princípios da proteção de dados. Sua inclusão

---

[559] ALHADEFF, J.; VAN ALSENOY, B.; DUMORTIER, J. The Accountability Principle in Data Protection Regulation: Origin, Development and Future Directions. *In*: GUAGNIN, D. *et al.* (org.). *Managing Privacy through Accountability*. London: Palgrave Macmillan UK, 2012. p. 49–82. Disponível em: https://doi.org/10.1057/9781137032225_4. Acesso em: 16 dez. 2020; e CASEY, B.; FARHANGI, A.; VOGL, R. Rethinking Explainable Machines: The GDPR's "Right to Explanation" Debate and the Rise of Algorithmic Audits in Enterprise. In: *Berkeley Technology Law Journal*, v. 34, p. 145-189, 2018. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3143325. Acesso em: 19 maio 2019.

[560] UNIÃO EUROPEIA. *WP 173 —* Opinion 3/2010 on the principle of accountability. [*S. l.*]: WP29, 2010. Disponível em: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2010/wp173_en.pdf. Acesso em: 21 jun. 2021.

[561] CASEY, B.; FARHANGI, A.; VOGL, R. Rethinking Explainable Machines: The GDPR's "Right to Explanation" Debate and the Rise of Algorithmic Audits in Enterprise. In: *Berkeley Technology Law Journal*, v. 34, p. 145-189, 2018. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3143325. Acesso em: 19 maio 2019.

[562] INTERNETLAB. O que está em jogo no debate sobre dados pessoais no Brasil? Relatório final sobre o debate público promovido pelo Ministério da Justiça sobre o Anteprojeto de Lei de Proteção de Dados Pessoais. São Paulo: InternetLab, 2016. Disponível em: https://www.internetlab.org.br/wp-content/uploads/2016/05/reporta_apl_dados_pessoais_final.pdf. Acesso em: 4 maio 2021.

deu-se posteriormente, na tramitação da câmara.

Embora o debate sobre a responsabilização dos agentes de tratamento perpassasse o tema da privacidade e proteção de dados desde as primeiras propostas de regulação, sua inclusão enquanto princípio reflete a maior preocupação com o estabelecimento de um regime mais claro de responsabilização dos agentes pela proteção de dados.

No contexto das decisões automatizadas, o princípio cumpre uma função primordial na garantia da proteção dos direitos dos titulares de dados. A previsão do princípio é um incentivo regulatório para que os agentes cumpram as obrigações da lei. Do ponto de vista da efetividade da autodeterminação informacional e dos direitos do titular dos dados pessoais, a explicação pode ser vista como uma prática de prestação de contas. Neste caso, prestação dos agentes de tratamento aos titulares de dados pessoais, ou em um segundo momento para o ente regulador. Apenas por essa prestação de contas e demonstração das responsabilidades é que o titular poderá exercer outros direitos, como oposição e de revisão das decisões automatizadas.

A proteção dos direitos sujeitos aos quais se referem os dados é um dos principais objetivos das regulações, de forma que um destinatário privilegiado desse dever de prestação de contas seja o titular de dados. Entretanto, é importante destacar que o dever de prestar contas extrapola o mero fornecimento de informações e garantias de acesso, como tratados no item 5.2, visto que a explicação pressupõe um processo relacional em que a efetiva compreensão do titular é essencial, conforme tratamos no item 4.2. Além disso, a previsão da responsabilização e da prestação de contas como princípio é um elemento que fortalece a possibilidade de uma efetiva reparação de danos causados a partir da análise dos riscos inerentes a determinadas atividades de tratamento de dados pessoais.

O sistema regulatório da proteção de dados se desenvolveu de forma a comportar diversos atores, para além das previsões individuais de privacidade, de forma que as obrigações de prestação de contas envolvem uma série de sujeitos aos quais os agentes de tratamento devem se reportar. Se, em primeiro lugar, destacamos os titulares de dados, em segundo lugar, podemos incluir como destinatários do dever de prestação de contas os reguladores.

Os relatórios de impacto e as auditorias previstas na LGPD poderão ser realizados a pedido da autoridade nacional de proteção de dados. O regimento interno

da ANPD, publicado em março de 2021, deu os primeiros passos para estabelecimento deste regime, com a criação da Coordenação Geral de Fiscalização, que, conforme disposição do art. 17 do regimento, possui a tarefa de fiscalizar a atuação dos agentes de tratamento públicos e privados.

Em terceiro lugar devemos considerar outros órgãos de estado que possuem competências concorrentes ou complementares à ANPD, como o poder judiciário e os órgãos de defesa do consumidor. A depender do processamento de dados e dos direitos afetados por uma decisão automática, haverá diferentes destinatários no poder público.

Ainda, conforme o regime de proteção de dados ganha capilaridade e relevância e a proteção de dados se consolida como direito mais reconhecido, a preocupação com o cumprimento das obrigações de proteção de dados torna-se uma preocupação mais comum no ambiente corporativo. As práticas de prestação de contas tornam-se elementos necessários para parceiros comerciais, clientes e investidores[563].

No contexto das decisões automatizadas, essa questão torna-se muito relevante, tendo em vista que o desenvolvimento de algoritmos para práticas comerciais envolve muitas vezes mais de um agente de tratamento, que pode assumir posições de controlador ou operador de dados. Se imaginássemos um algoritmo que automatiza determinada operação em um setor econômico, em casos em que esse algoritmo envolve o processamento de dados relacionados às pessoas naturais, logo nos deparamos com a obrigação de fornecer explicações relevantes sobre esta aplicação e de formas de demonstrar que tal desenvolvimento ocorreu em acordo com os princípios da LGPD.

Esse desenvolvimento pode ocorrer dentro de uma mesma organização. No entanto, pode ocorrer pela contratação de uma empresa especializada, de um fornecedor de serviço, que poderá realizar o desenvolvimento e a execução do processamento como operadora de dados. Nos casos em que mais de um agente de tratamento está envolvido, é possível imaginar que a obrigação de demonstrar o cumprimento da legislação e dados estará presente nas mais diferentes cláusulas

[563] RAJI, I. D. *et al.* Closing the AI accountability gap: defining an end-to-end framework for internal algorithmic auditing. *In*: FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency. New York, USA: Association for Computing Machinery, 2020. p. 33–44. Disponível em: https://doi.org/10.1145/3351095.3372873. Acesso em: 1 set. 2020.

contratuais e que em algumas delas as obrigações de prestação de contas e responsabilização estarão cuidadosamente elaboradas, incluindo a necessidade de desenvolvimento de instrumentos para tanto.

A LGPD estabeleceu a possibilidade da criação de selos e certificados, a exemplo da experiência europeia, que podem contribuir para a demonstração do cumprimento das obrigações de proteção de dados. Os mecanismos de certificação podem aumentar a transparência na relação das empresas com os titulares e com parceiros. Embora não sejam obrigatórios, a regulação europeia criou mecanismos efetivos para estimular sua adoção. O art. 83 do GDPR, por exemplo, estabelece que a adoção desses mecanismos pode influenciar a decisão e o arbitramento de sanções pelas autoridades supervisoras[564]. Certificações que possam garantir algoritmos mais confiáveis em termos de legalidade, de precisão e não discriminação podem contribuir para garantir maior segurança aos titulares e parceiros. Embora tal incentivo expresso não esteja presente na LGPD, é possível argumentar que a aderência a certificações e medidas de mitigação de riscos podem ser consideradas positivamente nas decisões administrativas do ordenamento brasileiro, devido a aplicação do princípio da prevenção[565].

Outro elemento importante nas obrigações de prestação de contas diz respeito à sociedade como um todo. Discutimos nos Capítulos 2 e 3 como as decisões automáticas podem afetar não apenas os direitos individuais, mas ter impactos sociais significativos. Dessa forma, os agentes de tratamento devem prestar contas à sociedade como um todo, inclusive para organizações da sociedade civil, que atuam para garantir valores importantes como a inclusão, a não discriminação e a diversidade.

As decisões automatizadas, portanto, podem contribuir para ganhos de produtividade e melhorias na qualidade de vários processos sociais. No entanto, tal qualidade e melhoria dependerá do estabelecimento de processos de prestação de contas dos vários agentes envolvidos. A frente abordaremos em mais detalhes alguns

---

564 EUROPEAN DATA PROTECTION BOARD. *Guidelines 1/2018 on certification and identifying certification criteria in accordance with Articles 42 and 43 of the Regulation — version adopted after public consultation*. [*S. l.*]: EDPB, 2018. Disponível em: https://edpb.europa.eu/our-work-tools/our-documents/leitlinien/guidelines-12018-certification-and-identifying-certification_en. Acesso em: 17 nov. 2020.

565 BIONI, B. R. Compreendendo o conceito de anonimização e dado anonimizado. *Revista do Advogado*, n. 144, nov. 2019. Disponível em: https://aplicacao.aasp.org.br/aasp/servicos/revista_advogado/paginaveis/144/20/index.html. Acesso em: 21 jul. 2021.

processos de prestação de contas, como relatórios de impacto e *explainability by design*.

5.1.4 Relatórios de impacto

Autores destacam que um dos maiores avanços das novas regulações de proteção de dados, entre elas a LGPD, diz respeito ao estabelecimento de obrigações mais fortes para os controladores, além do empoderamento dos titulares, algo que já vem se desenvolvendo desde gerações de regulações anteriores. Katerin Demetzu apresenta como os Data Processing Impact Assessment (DPIA) assumem a função de tornar o controlador responsável pela avaliação dos riscos de suas atividades.[566]

Antes de figurar como uma prática mandatória no regulamento, a Diretiva 95/46/CE da União Europeia trazia o conceito de *Privacy Impact Assessment* (PIA) como uma recomendação. É possível encontrar outras manifestações sobre os *Privacy Impact Assessment* (PIA) em orientações da autoridade australiana de proteção de dados em 2006[567] e na diretiva do governo canadense de 2010[568] .

Importantes trabalhos já apontavam como a exigência de DPIA poderia incentivar usos mais responsáveis dos dados pessoais, principalmente se tais documentos estivessem submetidos a supervisão de uma agência reguladora[569]. O conceito atual surge apenas no art. 35 do GDPR como mandatório para operações de alto risco. A opinião do WP29 definiu o DPIA como um instrumento que deve descrever o processamento de dados realizado, avaliar a sua necessidade, sua proporcionalidade e auxiliar na gestão dos riscos às liberdades e direitos dos titulares, indicando, para cada um deles, as medidas adequadas para seu enfrentamento. É uma forma de construir e demonstrar adequação ao princípio da *accountability*[570].

---

[566] DEMETZOU, K. Data Protection Impact Assessment: A tool for accountability and the unclarified concept of 'high risk' in the General Data Protection Regulation. *Computer Law & Security Review*, [*S. l.*], v. 35, n. 6, p. 105342, 2019. Disponível em: https://doi.org/10.1016/j.clsr.2019.105342. Acesso em: 21 jun. 2021.

[567] OAIC. *Privacy Impact Assessment Guide*. Sydney: OAIC, 2010. Disponível em: http://www.icb.org.au/out/?dlid=38156. Acesso em: 14 abr. 2021.

[568] TREASURY BOARD OF CANADA. *Directive on Privacy Impact Assessment*. 2010. Disponível em: https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=18308. Acesso em: 11 abr. 2021.

[569] QUELLE, C. *The Data Protection Impact Assessment, or*: How the General Data Protection Regulation May Still Come to Foster Ethically Responsible Data Processing. Rochester, NY: Social Science Research Network, 2015. SSRN Scholarly Paper. Disponível em: https://doi.org/10.2139/ssrn.2695398. Acesso em: 24 abr. 2021.

[570] EUROPEAN DATA PROTECTION BOARD. *Guidelines on Data Protection Impact Assessment (DPIA) (wp248rev.01)*. [*S. l.*]: EDPB, 2017. Disponível em: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611236. Acesso em: 17 nov. 2020.

Inserido num conjunto de elementos relacionados a um regime protetivo que estabelece a responsabilização dos agentes de tratamento, a obrigação de realização de um relatório de proteção de dados demanda destes uma postura proativa para conhecer a própria atividade. É uma forma de controle *ex ante* aos processamentos de dados onde se devem considerar todos os princípios da proteção de dados, por iniciativa do próprio agente. A obrigação de sua documentação é apontada como elemento central para a abordagem co-regulatória da proteção de dados, onde o poder público e os agentes privados atuam de forma conjunta para promover a proteção de dados e preservar o ambiente de inovação[571].

Encarar o relatório, contudo, como mera obrigação regulatória, ignora aspectos importantes sobre a avaliação de impacto. Como destaca Maria Cecília Gomes, os relatórios de impacto vão além de um documento de demonstração da conformidade à legislação de proteção de dados. É antes de tudo um instrumento de apoio das atividades de tratamento das organizações, como elemento essencial de sua governança e que deve focar na proteção de direitos dos titulares, e não na descrição da conformidade da organização[572].

A autora propõe uma segmentação analítica para a compreensão do que é e para que serve um relatório no contexto da LGPD. Em primeiro lugar, podemos considerar os Relatórios de Impacto à Proteção de Dados como um instrumento jurídico exigido pela legislação. Algumas autoridades de proteção de dados sugerem metodologias e modelos de construção desses relatórios. É importante destacar que sob esse ponto de vista, os relatórios se destinam a um agente em especial, qual seja, às autoridades de proteção de dados e outros órgãos da administração pública.

Em segundo lugar, podemos considerar os relatórios como uma análise, um processo de cognição. Não seria um mero registro, mas uma avaliação de processos e procedimentos. Sob essa dimensão ele torna-se central no contexto das decisões automatizadas. Esse processo de cognição deve resultar em um conhecimento capaz de orientar as ações do próprio agente de tratamento e, além disso, ser compartilhado com os titulares de dados ou sujeitos que serão impactados pelas decisões automatizadas, para que possam exercer ações realmente informadas.

---

[571] BINNS, R. Data Protection Impact Assessments: A Meta-Regulatory Approach. *International Data Privacy Law*, v. 7, n. 1, p. 22-35, 2017. Disponível em: https://papers.ssrn.com/abstract=2964242. Acesso em: 24 abr. 2021.

[572] GOMES, M. C. Relatório de impacto à proteção de dados. Uma breve análise da sua definição e papel na LGPD. *Revista do Advogado*, [*S. l.*], v. 39, n. 144, p. 174–183, 2019. p. 8.

A definição do relatório, presente no Inciso XVII do art. 5º da LGPD, apresenta ao menos duas modalidades de riscos que devem ser avaliados pelo relatório: os riscos às liberdades civis e aos direitos fundamentais. A interpretação sobre o que sejam direitos fundamentais e liberdades civis que podem ser impactadas por decisões automatizadas são amplas e dão margem a um debate à parte, mas cabe lembrar a discussão da ADIN nº 6.387/2020,[573] que se embasou no art 5° da Constituição Federal, em uma interpretação extensiva, para reconhecer a autodeterminação informacional e a proteção de dados pessoais como direito fundamental.

Uma vez que os relatórios de impacto permitem procedimentalizar a análise de riscos, é possível identificar ao menos dois sentidos para o termo risco no contexto da regulação e dos negócios. Uma delas diz respeito a uma definição vernacular. Risco, nesse sentido, significa a possibilidade da ocorrência de eventos indesejados no futuro[574]. Essa definição depende da definição do que seriam resultados desejados ou indesejados, e neste sentido, relaciona-se à valores compartilhados socialmente e no ordenamento jurídico. O estabelecimento de princípios para a proteção de dados, neste sentido, amplia o espectro de eventos que seriam considerados indesejados, ao estabelecer valores e orientações para a tutela de direitos subjetivos.

Outro sentido em que o termo risco é utilizado diz respeito a uma dimensão técnica e não necessariamente negativa. O risco é um elemento utilizado no cálculo para tomada de decisões, sobre a conveniência subjetiva de determinadas ações tomadas por agentes econômicos. As atividades econômicas implicam sempre em um risco, e quanto maior tal risco, de forma geral, maior poderá ser o retorno econômico de tal atividade.

Nas decisões automatizadas podemos identificar ao menos dois níveis de riscos. O primeiro deles diz respeito aos efeitos das decisões em si. Sejam tomadas por seres humanos ou automatizadas, decisões possuem o risco de resultar numa conclusão errônea. É possível que resulte na absolvição de um réu que seja culpado ou pode resultar na prisão de um réu inocente. Este último erro pode afetar um direito

---

[573] BRASIL. Supremo Tribunal Federal. *ADI nº 6.387/DF*. Rel. Min. Rosa Weber. Data de Publicação: 12/11/2020. Disponível em: http://portal.stf.jus.br/processos/detalhe.asp?incidente=5895165. Acesso em: 12 abr. 2021.

[574] 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. Disponível em: https://doi.org/10.1016/j.clsr.2017.12.003. Acesso em: 21 jun. 2021.

fundamental e causar um dano irreparável.

Já faz parte do conhecimento geral de que esse é um dos motivos pelos quais o processo penal se orienta por princípios como o da presunção de inocência, o devido processo legal e o direito ao contraditório. Espera-se que tais garantias possam minimizar o risco de tal tipo de erro. A aplicação desses princípios é uma escolha política dos ordenamentos jurídicos contemporâneos, que implica no aumento do risco do primeiro tipo de erro.

Tal escolha é comum em qualquer tipo de processo indutivo de conhecimento em que se busca afirmar a verdade de um fato e por esse motivo é importante estarmos atentos em relação a apelos tecnicistas que delegam as tecnologias à potencialidade de superar problemas morais complexos de nossa sociedade. O desenvolvimento de aplicações que possam afetar direitos ou causar impactos sociais deve orientar-se pelas mesmas preocupações.

Um algoritmo de análise antifraude, por exemplo, pode liberar uma compra para um fraudador. Por outro lado, pode negar um pedido de comprador de boa-fé. O desenvolvimento dessas aplicações permite diferentes níveis de precisão e erros e as aplicações estatísticas podem partir de diferentes estratégias. A partir de escolhas metodológicas os resultados podem privilegiar algum tipo de erro específico.

O segundo tipo de risco relacionado às decisões automatizadas diz respeito à escala de aplicação. Uma premissa equivocada ou um algoritmo enviesado pode reproduzir todo tipo de erro em escalas muito maiores do que determinadas decisões tomadas por humanos e em um curto espaço de tempo. E, nesse sentido, mesmo pequenos erros podem ter impactos significativos, por reproduzirem-se em escala. Mesmo que algoritmos sejam bem projetados e testados, livres de vieses, algumas aplicações disruptivas podem alterar comportamentos e causar impactos sociais significativos e imprevistos.

Logo, por essa razão, além de um documento legal e de uma avaliação de riscos, uma terceira forma de compreender os relatórios de impacto é considerá-los como um plano. As medidas para mitigação dos riscos e salvaguardas podem ser vistas como elementos da estrutura de governança dos agentes de tratamento que devem orientar-se pelos princípios da LGPD. Nesse quesito, quando agentes de tratamento identificam que determinada decisão é realizada com base em tratamento automatizado de dados, deve considerar os riscos de tais decisões, as possibilidades de falsos positivos e negativos, os possíveis vieses oriundos das escolhas

metodológicas ou das bases de dados. Além disso, deve considerar as obrigações de transparência e fornecer os elementos necessários para garantir aos sujeitos a possibilidade de compreender e, quando possível, defender-se dessa decisão. Em outras palavras, deve fornecer uma explicação sobre o algoritmo, que inclua os limites e riscos. Por consequência, o relatório de impacto direciona-se não apenas às autoridades de proteção de dados, mas aos titulares de dados e a sociedade civil como um todo.

O dever de avaliar riscos é extremamente relevante para as aplicações emergentes no campo da inteligência artificial. Neste sentido, a literatura do campo do *machine learning* tem enfrentado o problema e desenvolvendo metodologias para avaliação de impactos das aplicações. Essa avaliação de impacto extrapola a mera avaliação de risco de erros individuais. Há importantes discussões sobre o conceito de justiça e equidade entre a comunidade da ciência da computação e do direito, que inspiram modos de avaliar o potencial discriminatório e lesivo de aplicações de inteligência artificial, identificar impactos desiguais das aplicações para grupos sociais específicos. É importante destacar como processos discriminatórios contra determinados grupos podem ocorrer ainda que tais aplicações sejam “cegas” para essas categorias como raça ou gênero[575].

Nos últimos anos uma série de trabalhos têm apresentado diferentes metodologias para avaliação de impactos para decisões automatizadas como *Ethical Impact Assessment* (EIA)[576], o *Algorithm Impact Assessment* (AIA),[577] ou os *Human Rights and Ethical Impact Assessment* (HREIA)[578]. Tratam-se, cabe destacar, de esforços que vão além de problemas de nomenclatura. Essas propostas visam incorporar-se ao arcabouço dos DPIAs ou se apresentam como superações. No contexto norte americano, a discussão tem apresentado propostas relevantes diante

---

[575] MACCARTHY, M. Standards of Fairness for Disparate Impact Assessment of Big Data Algorithms. *Cumberland Law Review*, [*S. l.*], v. 48, n. 1, p. 67–148, 2017.

[576] WRIGHT, D.; MORDINI, E. Privacy and Ethical Impact Assessment. *In*: WRIGHT, D.; DE HERT, P. (org.). *Privacy Impact Assessment.* Dordrecht: Springer Netherlands, 2012. p. 397–418. Disponível em: https://doi.org/10.1007/978-94-007-2543-0_19. Acesso em: 24 abr. 2021.

[577] KAMINSKI, M. E.; MALGIERI, G. Algorithmic impact assessments under the GDPR: producing multi-layered explanations. *International Data Privacy Law*, [*S. l.*], n. ipaa020, 2020. Disponível em: https://doi.org/10.1093/idpl/ipaa020. Acesso em: 29 mar. 2021; e REISMAN, D. *et al. Algorithm Impact Assessment:* a practical framework for public agency accountability. [*S. l.*]: AI Now Institute, 2018. Disponível em: https://ainowinstitute.org/aiareport2018.pdf. Acesso em: 21 abr. 2021.

[578] MANTELERO, A. AI and Big Data: A blueprint for a human rights, social and ethical impact assessment. *Computer Law & Security Review,* [*S. l.*], v. 34, n. 4, p. 754–772, 2018. Disponível em: https://doi.org/10.1016/j.clsr.2018.05.017. Acesso em: 21 jun. 2021.

de diagnósticos similares dos problemas relacionados à inteligência artificial. É preciso considerar, por óbvio, as diferenças entre os níveis de regulação na proteção de dados e o próprio modelo e a tradição regulatória de cada contexto. Froomklin discute como a regulação no campo da proteção de dados pode se inspirar nas avaliações de impacto ambiental, propondo a realização de *Privacy Impact Notice*[579] (PIN). A proposta dos PIN diz respeito aos impactos da privacidade de aplicações que implicam em vigilância de massa. Segundo o autor, as avaliações de impacto ambiental, como os *Environmental Impact Statements* (EIS) tiveram uma efetividade em reduzir os riscos ambientais dos empreendimentos nos EUA. Uma avaliação sobre a privacidade poderia mitigar os riscos na esfera da proteção de dados. A proposta, contudo, permanece restrita ao conceito de privacidade.

Outras propostas visam superar o escopo das avaliações de impacto focadas na privacidade. O diagnóstico comum diz respeito a uma limitação da concepção que se origina dos tradicionais *Privacy Impact Assessment*. Caso sejam compreendidos de forma restrita ao conceito de proteção de dados pessoais, os relatórios podem observar e endereçar apenas os riscos relacionados à qualidade dos dados e a segurança das informações[580]. Essa abordagem restrita a proteção de dados pode não ser suficiente para dar conta dos riscos que emergem da inteligência artificial e dos usos de big data, de forma que o controle e a regulação das decisões automatizadas deve expandir-se para englobar uma dimensão de proteção de direitos como um todo[581].

Sonya Katial propõe uma avaliação também inspirada na regulação ambiental e no GDPR para a realização de um *Human Impact Statement* (HIS). A autora aponta as ameaças de práticas discriminatórias e aos direitos civis como os principais riscos envolvidos em aplicações de BIG data e de Inteligência artificial. Inspirada pelos DPIA, o contexto europeu, que na visão da autora inclui em seu escopo a avaliação ampla sobre os direitos dos titulares de dados. A proposta invoca o devido processo para a realização de decisões automatizadas e busca endereçar os riscos de decisões

---

[579] FROOMKIN, A. M. Regulating Mass Surveillance as Privacy Pollution: Learning from Environmental Impact Statements. *U. Ill. L. Rev.*, [*S. l.*], v. 2015, 2015. p. 1713.

[580] MANTELERO, A. AI and Big Data: A blueprint for a human rights, social and ethical impact assessment. *Computer Law & Security Review,* [*S. l.*], v. 34, n. 4, p. 754–772, 2018. Disponível em: https://doi.org/10.1016/j.clsr.2018.05.017. Acesso em: 21 jun. 2021.

[581] KATYAL, S. *Private Accountability in the Age of Artificial Intelligence*. Rochester, NY: Social Science Research Network, 2019. SSRN Scholarly Paper. Disponível em: https://papers.ssrn.com/abstract=3309397. Acesso em: 21 abr. 2021.

enviesadas e discriminatórias contra minorias sociais no processamento de dados.

Wright e Mondini propuseram um modelo de avaliação complementar aos tradicionais PIA, que poderiam fornecer aos stakeholders informações importantes para a tomada de decisões no desenvolvimento de aplicações baseadas no processamento de dados pessoais, chamado de *Ethical Impact Assessment* (EIA)[582]. A avaliação deveria ser levada em consideração para identificar e mitigar questões éticas. Os critérios para avaliação ética poderiam ser o impacto sobre a autonomia, à dignidade, ao consentimento informado, a justiça, a equidade, a coesão social, a segurança, a inclusão social, a não discriminação e a sustentabilidade.

Alguns autores sugerem, para as aplicações nas decisões automatizadas, a execução de *Algorithm Impact Assessment* (AIA)[583]. Uma abordagem que discute essas avaliações no contexto europeu pode contribuir para a compreensão dessas avaliações para a regulação brasileira. Na esteira da obrigação de realização de DPIA, Kaminski e Malgieri propõem que tais avaliações contenham as avaliações de riscos e estabeleçam um sistema que atenda as duas camadas de proteção de dados que podem ser depreendidas do regulamento: a) um modelo governança e salvaguarda dos riscos do processamento e b) a garantia dos direitos individuais dos titulares de dados[584].

Nesse sentido os autores propõe uma explicação multicamadas (*multilayered explanation*), com esforços multidisciplinares para refletir sobre aspectos técnicos e éticos, desde a definição do que seja um problema, das formas de medir os riscos e como os mitigar, além de estabelecer políticas de transparência que correspondam às exigências legais ou éticas.

Alessando Mantelero, propõe, a partir do exemplo dos *Human Rights Impact Assessments* (HRIA), adotados em um largo espectro de atividades econômicas, uma abordagem que permita extrapolar as avaliações focadas apenas nos aspectos relacionados à qualidade dos dados e a segurança das informações. HRIA

---

582 WRIGHT, D.; MORDINI, E. Privacy and Ethical Impact Assessment. *In*: WRIGHT, D.; DE HERT, P. (org.). *Privacy Impact Assessment.* Dordrecht: Springer Netherlands, 2012. p. 397–418. Disponível em: https://doi.org/10.1007/978-94-007-2543-0_19. Acesso em: 24 abr. 2021.

583 KAMINSKI, M. E.; MALGIERI, G. Algorithmic impact assessments under the GDPR: producing multi-layered explanations. *International Data Privacy Law*, [*S. l.*], n. ipaa020, 2020. Disponível em: https://doi.org/10.1093/idpl/ipaa020. Acesso em: 29 mar. 2021; REISMAN, D. *et al. Algorithm Impact Assessment:* a practical framework for public agency accountability. [*S. l.*]: AI Now Institute, 2018. Disponível em: https://ainowinstitute.org/aiareport2018.pdf. Acesso em: 21 abr. 2021.

584 KAMINSKI, M. E.; MALGIERI, G. Algorithmic impact assessments under the GDPR: producing multi-layered explanations. *International Data Privacy Law*, [*S. l.*], n. ipaa020, 2020. Disponível em: https://doi.org/10.1093/idpl/ipaa020. Acesso em: 29 mar. 2021.

tradicionalmente adotado para avaliação de tratados de comércio internacional fornece um modelo promissor para a avaliação de aplicações de inteligência artificial, ao basear-se em um ordenamento uniforme e validado para diversos contextos nacionais. As regulações de direitos humanos permitem avaliar impactos e riscos aos direitos das comunidades tradicionais, às minorias étnico raciais, às mulheres, ao desenvolvimento econômico de populações vulneráveis, à democracia, ao emprego, ao meio ambiente, ao exercício das liberdades individuais[585].

O autor propõe que se realizem o que poderia ser chamado de *Human Rights and Ethical Impact Assessment* (HREIA) para as aplicações de inteligência artificial que possam impactar direitos, entre elas as utilizadas em decisões automatizadas e oferece uma metodologia para diversas modalidades de relatório, *ex ante* ou *ex post,* incorporando mais camadas de análise nas avaliações de impacto, compostas pela dimensão dos direitos humanos e de avaliações éticas, como forma de ir além de legislações locais de proteção de dados[586].

As são extremamente relevantes ao demonstrar como os instrumentos já estabelecidos na regulação, seja sobre a proteção de dados, sejam as regulações de direitos humanos ou regulações setoriais sobre igualdade, direito e liberdades civis ou direitos fundamentais podem servir de ferramenta efetiva para enfrentar os riscos e impactos das decisões automatizadas e das atividades baseadas em inteligência artificial, a despeito das incertezas significativas sobre o conceito de risco e os escopos de proteção da proteção de dados que já discutimos neste trabalho.

Como se pode notar, o conceito de risco surge como um elemento central para a compreensão do escopo da avaliação e da sua obrigatoriedade tanto no contexto europeu quanto no brasileiro, embora haja diferenças significativas. A GDPR estipula a obrigatoriedade para aplicações de alto risco, o que, de início, já suscita a questão sobre a definição do que seja alto risco. Os *Recitals* 75 e 76 apontam que tal avaliação deve se balizar pela severidade e a probabilidade (*severity and likelyhood*) de que a operação cause impactos negativos aos direitos e liberdades fundamentais.

A severidade e a probabilidade de impactos fornecem critérios ainda muito amplos para determinar com exatidão quando os DPIA seriam necessários. Podemos

---

585 WALKER, S. M. *The future of human rights impact assessments of trade agreements*. [*S. l.*]: [*s. n.*], 2009.

586 MANTELERO, A. AI and Big Data: A blueprint for a human rights, social and ethical impact assessment. *Computer Law & Security Review,* [S. l.], v. 34, n. 4, p. 754–772, 2018. Disponível em: https://doi.org/10.1016/j.clsr.2018.05.017. Acesso em: 21 jun. 2021.

deduzir uma avaliação de 3 níveis antes de determinar sua relação: a) a avaliação da existência de riscos, b) a avaliação da severidade dos eventos, c) a probabilidade de que eles aconteçam[587]. As *Guidelines on Data Protection Impact Assessment,* de 2017, produzida pelo EDPB, lista uma série de aplicações para as quais os DPIA seriam mandatórios. O rol tem exemplos mais precisos, embora não exaustivos, para a realização do relatório. Entre elas, inclusive, encontra-se o tratamento de dados que envolva decisões automatizadas a partir de perfilamento de indivíduos[588].

A obrigação no contexto brasileiro é menos precisa. A LGPD traz o "Relatório de Impacto a Proteção de Dados" em diversos artigos. Sua definição mais precisa aparece no inciso XVII do Art. 5° e no parágrafo único do Art. 38, segundo os quais o relatório deve avaliar os riscos às liberdades civis e aos direitos fundamentais, deve conter a descrição dos dados coletados, a metodologia da coleta e da segurança do processamento e por fim deve conter uma análise do controlador com relação às salvaguardas e medidas de segurança para mitigar tais riscos. De acordo com o regimento interno da ANPD, caberá à Coordenação-Geral de Supervisão requisitar a apresentação do Relatório de Impactos à Proteção de Dados aos agentes de tratamento.

Em nenhum momento a LGPD restringe a necessidade de execução do relatório a eventualidade de tratamento de dados de alto risco, tampouco determina hipóteses para sua obrigatoriedade[589]. As previsões da lei delegam à autoridade nacional uma grande discricionariedade para exigir a sua realização.

E cabe destacar, além disso, conforme destacado no item 5.1, que embora o relatório aparece vinculado a ações dos controladores de dados, nos contextos de decisões automatizadas, a obrigação de avaliar os riscos e mitigá-los pode incluir também operadores de dados pessoais, na medida em que são os responsáveis pelas

---

[587] DEMETZOU, K. Data Protection Impact Assessment: A tool for accountability and the unclarified concept of 'high risk' in the General Data Protection Regulation. *Computer Law & Security Review*, [*S. l.*], v. 35, n. 6, p. 105342, 2019. Disponível em: https://doi.org/10.1016/j.clsr.2019.105342. Acesso em: 21 jun. 2021.

[588] EUROPEAN DATA PROTECTION BOARD. *Guidelines on Data Protection Impact Assessment (DPIA) (wp248rev.01)*. [*S. l.*]: EDPB, 2017. Disponível em: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611236. Acesso em: 17 nov. 2020.

[589] "Art. 38. [...] Parágrafo único. Observado o disposto no caput deste artigo, o relatório deverá conter, no mínimo, a descrição dos tipos de dados coletados, a metodologia utilizada para a coleta e para a garantia da segurança das informações e a análise do controlador com relação a medidas, salvaguardas e mecanismos de mitigação de risco adotados." (BRASIL. *Lei nº 13.709, de 14 de agosto de 2018*. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF: Presidência da República, 14 ago. 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015- 2018/2018/lei/L13709.htm. Acesso em: 07. nov. 2020).

aplicações e pela decisão. Mesmo que um operador realize tratamento em nome de outrem, na medida em que é responsável pela lógica do processamento e da integridade do algoritmo, tais operadores devem garantir meios de explicação e, quando possível, meios de revisão das decisões.

Uma das possíveis consequências do desenvolvimento de relatórios de impacto é uma análise da forma como os dados são tratados, o que permite implementar modificações no *design* dos processos, das aplicações e dos sistemas desde a sua concepção. No caso de processos automatizados, relatórios podem permitir que estes sejam desenvolvidos de tal forma que a explicação sobre o seu funcionamento seja parte inerente do sistema, conceito que vem sendo chamado de *explainability by design*, que será explorado em seguida.

5.1.5 Os desenvolvedores: *explainability by design*

No item 4.2, discutimos sobre a natureza das explicações e de como estas constituem uma simplificação da realidade. Apresentamos, também, como algumas aplicações apresentam limites para o próprio entendimento dos processos lógicos internos aos modelos de *machine learning*. Já discutimos as limitações cognitivas em função de características culturais ou socioeconômicas, bem como as questões relativas a segredos de negócio e propriedade intelectual.

Nesta seção iremos discutir sobre o desenvolvimento de modelos explicáveis, principalmente no campo da inteligência artificial, apresentando formas de cumprir com as obrigações de transparência e *accountability* desde a concepção de aplicações ou sistemas. Em outras palavras, em relação a desenvolver *explicability by design*, em diálogo com as metodologias já consolidadas de *privacy by design* implicitamente prevista na LGPD.

O conceito de *privacy by design* surgiu nos anos 1990 com o objetivo de promover maior confiança nas plataformas de informação e comunicação. Aos poucos tal conceito foi se inserindo nas referências de autoridades nacionais de proteção de dados pessoais desde o Canadá até a Europa. Além disso, o conceito expande sua abrangência para sistemas organizacionais em geral e outras arquiteturas tecnológicas conforme debates sobre a tecnologias e a proteção de dados

amadureciam em diversos países[590].

Uma série de trabalhos apontava que problemas em algumas aplicações e nos algoritmos decorriam muitas vezes de problemas na própria concepção de projeto. Redes sociais podem ser usadas como exemplos desse fenômeno. Mulligan e King apresentaram, em um trabalho de 2012, como se podia observar um descompasso entre o *design* das aplicações e a privacidade dos usuários[591], principalmente nas aplicações de redes sociais. Esse descompasso se manifestava no sentimento dos usuários das redes sentindo-se ao mesmo tempo *spamed*, ou seja, recebendo muitos conteúdos indesejados e também *stalked*, visto que não tinham segurança sobre quem teria acesso a suas informações pessoais. O problema principal na visão das autoras estava relacionado ao descumprimento das expectativas dos usuários decorrente de uma concepção falha sobre a privacidade, entendida apenas como uma questão técnica e não contextual.

A teoria da privacidade que deveria influenciar o *design* dessas aplicações e que estas deveriam levar em consideração a integridade dos fluxos informacionais e as expectativas dos usuários. A exposição de informações é o próprio objeto das relações entre plataformas e usuários das redes sociais, é o motivo para os quais as pessoas utilizam as plataformas. No entanto, as aplicações nem sempre continham mecanismos efetivos para que os usuários utilizassem sua capacidade de controle sobre os dados pessoais.

Nesse sentido, argumentava-se que os FIPs não eram suficientes para garantir a privacidade contextual e a garantia de que as expectativas dos usuários seriam atendidas. Para ir além das obrigações legais, o *design* pode fazer com que a tecnologia seja uma ferramenta para aumentar o controle dos titulares sobre seus dados.

O termo foi incorporado às legislações e é expressamente adotado pelo

---

[590] CAVOUKIAN, A. Privacy by Design: Origins, meaning, and prospects for assuring privacy and trust in the information era. *In*: YEE, G. O. M. *Privacy Protection Measures and Technologies in Business Organizations*: Aspects and Standards, [*S. l.*], IGI Global, 2012. p. 170–208. Disponível em: https://doi.org/10.4018/978-1-61350-501-4.ch007. Acesso em: 21 jun. 2021.

[591] MULLIGAN, D. K.; KING, J. *Bridging the Gap between Privacy and Design*. Rochester, NY: Social Science Research Network, 2012. SSRN Scholarly Paper. Disponível em: https://papers.ssrn.com/abstract=2070401. Acesso em: 23 set. 2020.

GDPR[592]. Na lei brasileira, o conceito aparece no art. 46[593], que determina adoção de medidas técnicas e administrativas desde a concepção de um produto ou serviço. Em razão dessa obrigação, alguns doutrinadores e atores do mercado têm defendido a adoção da metodologia de *explainability by design*, que determina a necessidade de incorporação, desde a concepção dos produtos e serviços, até à capacidade para garantir o direito à explicação das decisões tomadas por sistemas automatizados.

Se a LGPD garante um direito a informações significativas sobre o processamento desses dados, o mero fornecimento dessas informações pode não contribuir para o objetivo principal do direito à explicação, que é garantir aos titulares ou usuários de sistemas automatizadas de tomadas de decisão a capacidade de questionar e desafiar tais decisões de forma efetiva[594].

Os regulamentos de proteção de dados estabelecem critérios claros e precisos em vários institutos, como na forma do consentimento (livre e informado), e definem formas de exercer a transparência e o próprio conceito de *privacy by design* fornecem um corpo cogente de normas que garantem maior controle sobre os dados pessoais e uma maneira de superar as assimetrias informacionais. Mas o direito à explicação encontra ainda pouca formulação legal sobre quais parâmetros deveriam ser utilizados para a avaliação de sua efetividade. A ausência dessa previsão é o que motiva autores a afirmar a inexistência desse direito[595]. No entanto, uma série de

---

[592] "Art. 25. Data protection by design and by default. (1) Taking into account the state of the art, the cost of implementation and the nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for rights and freedoms of natural persons posed by the processing, the controller shall, both at the time of the determination of the means for processing and at the time of the processing itself, implement appropriate technical and organisational measures, such as pseudonymisation, which are designed to implement data-protection principles, such as data minimisation, in an effective manner and to integrate the necessary safeguards into the processing in order to meet the requirements of this Regulation and protect the rights of data subjects." (UNIÃO EUROPEIA. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance)*. Official Journal of the European Union, 4 maio 2016. Disponível em; https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng. Acesso em: 27 ago. 2021).

[593] "Art. 46. Os agentes de tratamento devem adotar medidas de segurança, técnicas e administrativas aptas a proteger os dados pessoais de acessos não autorizados e de situações acidentais ou ilícitas de destruição, perda, alteração, comunicação ou qualquer forma de tratamento inadequado ou ilícito. [...] § 2º As medidas de que trata o caput deste artigo deverão ser observadas desde a fase de concepção do produto ou do serviço até a sua execução." (BRASIL*. Lei nº 13.709, de 14 de agosto de 2018*. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF: Presidência da República, 14 ago. 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015- 2018/2018/lei/L13709.htm. Acesso em: 07. nov. 2020).

[594] DOSHI-VELEZ, F.; KIM, B. Towards A Rigorous Science of Interpretable Machine Learning. *ArXiv [cs, stat*], [*S. l.*], 2017. Disponível em: http://arxiv.org/abs/1702.08608. Acesso em: 30 jun. 2020.

[595] WACHTER, S.; MITTELSTADT, B.; FLORIDI, L. *Why a Right to Explanation of Automated Decision-*

trabalhos tem apontado caminhos para implementação e avaliação de explicações, além disso, órgãos regulatórios têm incentivado o desenvolvimento da explicabilidade de sistemas.

Nas seções anteriores discutimos como os direitos sobre a propriedade intelectual e o segredo comercial representam limites ao exercício da explicação. As limitações em função de segredos de negócio são um problema de que tratamos ao descrevermos a possibilidade de as autoridades nacionais realizarem auditorias nos algoritmos e os deveres de prestação de contas e responsabilização. Por conta dessas obrigações, como demonstramos, as organizações privadas e públicas iniciam processos para tornar os procedimentos mais transparentes e em alguns casos, mais explicáveis, principalmente nos casos de decisões automatizadas ou que possam causar impactos nos titulares de dados.

Relacionado ao tema da inteligência artificial, o campo de estudo conhecido como Explainable Artificial Intelligence (XAI) tem atraído cada vez mais atenção num crescente número de publicações e eventos mundo afora. Um levantamento bibliográfico sobre o tema destaca que, embora o termo seja recente – foi utilizado pela primeira vez em 2004 —, o problema já era conhecido pelo menos desde a década de 1970 pela comunidade de sistemas especializados[596].

O estudo analisou as conferências e publicações mais relevantes sobre o tema e identificou dois grandes atores no desenvolvimento das pesquisas. De um lado, os acadêmicos desenvolvendo pesquisas no campo que ficou conhecido pelo acrônimo FAT, dos termos anglo saxões *fairness, accountability e transparency.* O outro grupo importante é formado pelos trabalhos de pesquisadores civis e militares sob financiamento da agência estadunidense para pesquisas avançadas em defesa, a DARPA.

Mais recentemente, os setores industriais também iniciaram projetos de desenvolvimento de modelos explicáveis de inteligência artificial, pensando na relação entre os sistemas e os seus usuários finais, com o objetivo principal de aumentar a usabilidade e a confiança em seus produtos.

Making Does Not Exist in the General Data Protection Regulation. Rochester, NY: Social Science Research Network, 2016. Disponível em: https://papers.ssrn.com/abstract=2903469. Acesso em: 27 maio. 2020.

[596] ADADI, A.; BERRADA, M. Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *IEEE Access*, [*S. l.*], v. 6, p. 52138–52160, 2018. Disponível em: https://doi.org/10.1109/ACCESS.2018.2870052. Acesso em: 21 jun. 2021.

O campo de pesquisa aparecia inicialmente relacionado ao enfrentamento da opacidade dos modelos *black box*, onde a explicabilidade e a interpretabilidade apareciam muitas vezes como sinônimos. As motivações para a necessidade da explicabilidade eram em sua maioria relacionadas à justificação de decisões tomadas pelos sistemas para melhorar o controle sobre os sistemas, permitir sua melhoria ou ajudar a interpretar resultados em contexto de descobertas científicas ainda em estágio exploratório. Não restam dúvidas, contudo, que a consolidação do GDPR como um marco e contribuição importante na maior atenção dada ao tema.

Trabalhos mais recentes vêm consolidando metodologias para desenvolver sistemas interpretáveis em diversos tipos de aplicações. Há propostas para o desenvolvimento de modelos interpretáveis, chamados de w*hite box*, em oposição aos modelos *black box*[597]. Há, além disso, formas de interpretabilidade de modelos mais opacos, os quais podem ser objeto de interpretação a partir da criação de outros modelos capazes de fornecer informações relevantes que possam ser interpretadas e utilizadas em uma explicação.

Alguns trabalhos mais recentes tentam avançar em formas de se avaliar a interpretabilidade dos sistemas. Doshi-Velez e Been Kim propuseram uma taxonomia e formas de avaliação da interpretabilidade de sistemas de inteligência artificial[598]. A primeira forma de avaliar a explicação seria o que denominaram de *application grounded*, medindo a explicação no contexto da sua aplicação, em contextos específicos e com os *experts* que utilizam os modelos de *machine learning*. São avaliações de tarefas reais e feitas a partir da análise de pessoas reais no contexto de uso das aplicações e utilizando metodologias qualitativas e quantitativas.

A segunda forma de avaliação identificada pelos autores seria *human grounded*. Para esse tipo de avaliação, é preciso realizar testes simples que mimetizem as operações realmente realizadas pelos sistemas. Esses devem ser realizados com seres humanos vistos como o público-alvo da aplicação. Tais modos de avaliação são úteis quando precisam ser conduzidas em condições limitadas, como pouco tempo ou quando se busca medir a qualidade da explicação sob determinados aspectos isolados. Em termos gerais, são metodologias que utilizam pessoas reais e tarefas simplificadas de forma quantitativa ou qualitativa.

---

[597] MOLNAR, C. *Interpretable Machine Learning*. [*S. l.*]: [*s. n.*], 2020. *E-book*.
[598] DOSHI-VELEZ, F.; KIM, B. Towards A Rigorous Science of Interpretable Machine Learning. *ArXiv [cs, stat*], [*S. l.*], 2017. Disponível em: http://arxiv.org/abs/1702.08608. Acesso em: 30 jun. 2020.

A última forma de avaliação da explicação seriam *funcionally grounded*, a qual utiliza definições formais de interpretabilidade como representação da qualidade da explicação e realiza testes automatizadas medindo a eficácia em função desses parâmetros previamente definidos. Recomendado para aplicações onde não há recursos disponíveis para testes com pessoas, quando tais testes são antiéticos, onde ainda não há validação da metodologia ou em casos em que os testes já foram realizados suficientemente em outros contextos. São testes automáticos de tarefas simplificadas, meramente quantitativas e indiretas.

Para escolher a melhor forma de avaliar a explicação, os autores recomendam então que se analisem ao menos duas dimensões. Em primeiro lugar, as características e implicações da tarefa realizada pelo sistema, de forma que se tenha uma visão sobre a necessidade da explicação: se é preciso de uma visão global ou local do seu funcionamento; se há segmentações do problema; se há características específicas em alguma dessas partes e qual é a *expertise* dos usuários.

Em segundo lugar, deve-se analisar as características e implicações relativas ao método necessário para uma explicação eficiente: quais as peças cognitivas necessárias para a explicação e qual a sua natureza; se seria uma lista ou um número; quais são os elementos necessários para a explicação e qual a relação de composição entre eles; se há causalidade ou hierarquia e qual a relação entre os pedaços cognitivos da explicação.

No contexto da interação entre humanos e sistemas automáticos, explicações devem ser entendidas como representações de determinadas informações que, após processadas por um receptor, permitem a esse compreender determinado fenômeno. Convém apontar que informação diz respeito a aspectos factuais e materiais de uma realidade. Já representações podem ocorrer de diversas formas (gráficas, linguísticas etc). A informação é independente do agente, a representação, contudo, é direcionada.

Outro ponto importante ao pensarmos no desenvolvimento de sistemas explicáveis é que as explicações só se realizam após o processamento das informações por aquele que as recebe por meio de uma representação. Portanto, não basta que uma representação seja processável em abstrato, ela precisa ser processável em um contexto, de forma que seja compreendida pelo receptor.

Portanto, uma categoria central para a concepção de sistemas explicáveis é a definição de um grupo alvo para a explicação. Embora não seja possível definir

precisamente o que seja um grupo, explicações que não tenham intenção de realizar generalizações e abstrações são pouco úteis a menos que pensemos em sistemas completamente artesanais. Esse grupo alvo pode ser pensado a partir de agentes representativos que possuem determinadas características comuns.

Por último, é preciso pensar em formas de implementar a explicação, que podem ser geradas por outro sistema, pelo próprio sistema enquanto é utilizado ou por um humano especializado. Estudos de interação humano-máquina, pela psicologia cognitiva e os estudos de geração de linguagem natural são as disciplinas mais envolvidas nesse campo. Com a disseminação de atividades automatizadas e da utilização de dispositivos computacionais, as ciências sobre as interações entre os usuários e os sistemas têm ganhado cada vez mais espaço.

Alguns estudos têm apresentado modelos de explicações dos sistemas partindo de investigações sobre as possíveis necessidades dos usuários. A definição desse usuário é de extrema importância e pode decorrer de estudos empíricos ou de estudos de cenários. Ao focar nas necessidades dos usuários, destaca-se a diferença entre a interpretabilidade e a explicação: a primeira é uma característica dos sistemas e a segunda é uma ação intencional. Por isso, a explicação não se refere necessariamente ao funcionamento do sistema, mas oferece informações significativas e úteis para determinado interlocutor. Essas explicações poderiam se situar entre categorias comuns da literatura de *human computer interaction* (HCI): confiança, *human-likeness*, justificação adequada, e compreensibilidade[599]. Outros trabalhos apontam como se utilizar metodologias empíricas para identificar as melhores explicações e destacam a importância do desenvolvimento da interdisciplinaridade e da sensibilização da comunidade de desenvolvedores e cientistas de dados[600] .

Um relatório produzido pelo ICO e pelo The Alan Turing Institute resume uma série de recomendações de práticas para transparência e *accountability* para modelos baseados em inteligência artificial e apresenta parâmetros para tornar explicações

[599] EHSAN, U.; RIEDL, M. O. *On Design and Evaluation of Human-centered Explainable AI systems*. Glasgow, 2019. Disponível em: https://www.cc.gatech.edu/~riedl/pubs/ehsan-chi-hcml19.pdf. Acesso em: 22 set. 2020.

[600] LIAO, Q. V.; GRUEN, D.; MILLER, S. Questioning the AI: Informing Design Practices for Explainable AI User Experiences. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems,* [*S. l.*], p. 1–15, abr. 2020. Disponível em: https://doi.org/10.1145/3313831.3376590. Acesso em; 21 jun. 2021.

compreensíveis para as pessoas afetadas por decisões automatizadas[601]. O documento parte da análise de uma série de trabalhos sobre explicação e propõe uma abordagem voltada à implementação adequada de sistemas de inteligência artificial. Apesar de não ser um conjunto de normas legais, o documento produzido por uma autoridade de proteção de dados apresenta um quadro completo do que seria necessário levar em consideração na concepção de produtos ou serviços para que estes possam ser explicáveis.

O relatório identificou uma série de tipos de explicação que podem ser classificadas de maneiras diferentes. A primeira forma de classificação diz respeito ao objeto da explicação, que pode recair sobre o processo de um sistema, ou seja, suas características, a governança e seu *design* ou pode recair sobre os resultados desse sistema e as razões para uma razão específica, que corresponde a oposição tratada anteriormente entre explicações globais e locais e as necessidades de se conhecer uma decisão específica ou a lógica do sistema que toma decisões.

Além dessa primeira divisão, o relatório identificou 6 tipos diferentes de explicação:

i) Explicação da racionalidade: esse tipo de explicação se debruça sobre o porquê de uma decisão automática e permite o seu questionamento ou a modificação. Para que o sujeito ao qual se comunica compreenda a racionalidade de um sistema é preciso demonstrar como este chegou àquela conclusão, quais os parâmetros utilizados, quais são os mais relevantes, o quão confiáveis são eles. É preciso demonstrar como se pode embasar a decisão e como o sistema foi pensado levando em consideração a realidade das pessoas afetadas. Esse tipo de explicação pode estar tanto no processo quanto no resultado;
ii) Explicação da responsabilidade: aqui deve-se demonstrar quem é o responsável e como se pode rever uma decisão. O objetivo é permitir o questionamento e esclarecimento da decisão tomada. Por isso deve conter os sujeitos responsáveis pelo desenvolvimento e definições do sistema desde

[601] INFORMATION COMMISSIONER'S OFFICE; THE ALAN TURING INSTITUTE. ICO and Turing consultation on Explaining AI decisions guidance. *ICO*, 24 jan. 2020. Disponível em: https://ico.org.uk/about-the-ico/ico-and-stakeholder-consultations/ico-and-the-turing-consultation-on-explaining-ai-decisions-guidance/. Acesso em: 16 abr. 2021.

sua concepção e os mecanismos pelos quais é possível exercer controle sobre cada sujeito. É importante que constem as regras e processos internos da organização em cada estágio das decisões, incluindo fornecedores ou desenvolvedores. Por essas características é uma explicação baseada no processo, embora possa conter alguns elementos relacionados ao resultado. E conforme se apresentem os responsáveis, deve conter informações claras identificando um canal de comunicação e explicando como se pode utilizá-lo para revisão das decisões;

iii) Explicação dos dados: nessa explicação deve-se fornecer as informações claras sobre quais dados são utilizados, quais são suas fontes e as bases legais para sua utilização. Nesse sentido, permite o questionamento e a confiança nos resultados e no algoritmo. Por isso, deve fornecer qual dado se utilizou para treinar o modelo e quais foram os dados utilizados numa decisão específica. Ainda deve informar a metodologia e as salvaguardas para coleta e utilização desses dados, quais as precauções foram tomadas para avaliar a integridade dos dados e quais medidas foram tomadas para evitar potenciais vieses ou discriminações. Pode ser baseada no processo ou no resultado;

iv) *Fairness explanation*: a explicação de *fairness* consiste em demonstrar ao usuário as medidas que foram tomadas e os cuidados de monitoramento para garantir que essa decisão seja equilibrada e livre de viés. Essa explicação fornece confiança ao modelo e permite o questionamento das decisões. É preciso explicar se os dados são representativos, relevantes, precisos, imparciais, suficientes e atualizados em todas as etapas. Deve focar nas medidas e testes realizados para garantir a equidade das decisões. Pode focar no processo ou no resultado, mas principalmente no processo, visto que é preciso demonstrar que o próprio *design* é adequado, incluindo mitigação de riscos de viés decorrente do contexto de coleta, na escolha dos parâmetros e métricas ou da incidência de variáveis externas que podem determinar o resultado. É preciso demonstrar que o resultado é adequado explicitando os critérios utilizados para definição dos seus valores. Além disso, é preciso explicar como se avaliaram os possíveis impactos dessas decisões de *design*. É preciso demonstrar a adequação de sua implementação por pessoas treinadas, responsáveis e íntegras, conscientes das limitações do sistema;

v) Explicação de performance e segurança: a explicação de segurança auxilia

os usuários a compreenderem as medidas tomadas para garantir que o sistema é preciso, confiável, seguro e robusto. Pode-se comparar o desempenho dessas aplicações com a atuação de seres humanos realizando a mesma tarefa. O objetivo principal desse tipo de explicação é confirmar a integridade do sistema e permitir que os afetados possam questionar as decisões e por isso pode estar baseado no processo ou no resultado. Para isso deve informar o grau precisão (ou proporção de resultados corretos), como essa métrica foi calculada, a escolha dessa metodologia e qual o processo de validação externa utilizado. Deve ainda informar o método de monitoramento do sistema e quais são as medidas de confiança (*reliability*) utilizadas para avaliar se o sistema é capaz de realizar as tarefas para o qual foi programado, informando quais foram os procedimentos de análise formal dos códigos e das especificações do sistema. Essa explicação deve fornecer informações sobre a capacidade de a arquitetura do sistema estar livre de danos ou interferências externas. É necessário fornecer informações sobre a robustez do sistema, ou como este responde a situações não programadas ou adversas;

vi) Explicação e impacto: as explicações de impacto envolvem demonstrar como foram avaliados os efeitos das aplicações nas pessoas afetadas e na sociedade como um todo. Tem o propósito de fornecer aos indivíduos formas de agir e controlar suas ações e fornecer maior confiança de que os riscos foram sopesados antes da implementação do sistema. É preciso demonstrar como os riscos e consequências foram avaliados e quais medidas foram tomadas para minimizar esses riscos. Além disso, é importante demonstrar como esse impacto é constantemente monitorado. Embora sejam focadas em características gerais do sistema, é desejado que se demonstre como foram analisados os impactos para cada decisão.

O documento ainda elenca 5 fatores contextuais que podem afetar o tipo de explicação necessária como a) os setor de aplicação do modelo de I.A., por exemplo, setor criminal ou de seguros; b) o impacto das decisões na vida, no patrimônio, nos direitos ou nas liberdades; c) a forma como os dados utilizados afetam os indivíduos; d) a forma como a urgência da decisão pode afetar a percepção do indivíduo sobre a necessidade de um procedimento e sua proporcionalidade; e) a audiência para a qual

se destina a explicação a depender do grau de expertise, o quanto essa audiência é afetada etc.

O relatório apresenta então um sumário de tarefas para desenvolvimento de inteligência artificial explicável que engloba várias etapas a partir da concepção até a implementação dos sistemas:

1) A partir de princípios e dimensões éticas, é preciso delimitar quais são as explicações necessárias. Partir do domínio da aplicação e da experiência dos usuários para definir quais são as principais necessidades. É crucial que se realize uma análise de impactos individuais e sociais dessas aplicações para que as explicações sejam relevantes e signifiquem o incremento na transparência da empresa. É importante destacar que há um cálculo de prioridades, visto que nem toda explicação é relevante e os interesses podem variar de acordo com cada usuário de forma que se deve pensar em formas de garantir contextualmente a melhor explicação;
2) Coletar os dados de forma adequada e pensando em explicabilidade. A forma como dados são coletados e pré-processados impactam na qualidade das informações. Dessa forma, é importante levar em consideração as fontes dos dados, sua representatividade e as salvaguardas, internas ou externas, em relação a controladores, operadores ou co-controladores de dados;
3) Construir o sistema pensando em sua capacidade de fornecer explicações, pensar em modelos suplementares quando necessário nos casos em que o modelo for opaco por natureza;
4) Traduzir a racionalidade do sistema para razões compreensíveis e úteis para o receptor de forma que possam ser utilizadas. Dessa forma, é importante explicitar os pressupostos matemáticos usados para cada conclusão ou na produção de variáveis, indicadores ou métricas de forma compreensíveis para não técnicos;
5) Preparar os implementadores para utilizar e interpretar o modelo corretamente e a reconhecer seus limites e usos específicos;
6) Considerar a forma da explicação de acordo com o contexto da explicação em função do público-alvo e das características das interações entre os usuários e o sistema, podendo incluir textos, gráficos, etc.

Há argumentos de que a incorporação de explicabilidade pode significar um aumento nos custos de desenvolvimento e implementação de tecnologias de inteligência artificial, principalmente no contexto de empresas enxutas e modelos de gestão ágil focados em resultados rápidos. Essas metodologias, contudo, podem representar vantagens e oportunidades para melhor gestão de riscos e incremento da confiança dos usuários na segurança das aplicações. O desenvolvimento de plataformas realmente transparentes que empoderem os usuários ou pessoas afetadas por suas decisões pode representar uma vantagem competitiva[602].

Além disso, há certas aplicações de alto risco para as quais avaliações de impacto e salvaguardas devem ser exigidas, mesmo que signifique mais esforços ao desenvolvimento tecnológico[603], visto que seus impactos representaram custos sociais significativos.

5.1.6 Auditoria

Anteriormente, discutimos a importância do direito à explicação para garantir direitos dos titulares nas decisões automatizadas. Os principais argumentos em defesa desse direito no ordenamento europeu, mesmo que não tenha sua previsão expressa no GDPR, está associado ao entendimento de que regulamento de proteção de dados institui um regime de *accountability* geral no tratamento de dados pessoais, incluindo o processamento para fins de decisões automáticas, do qual a explicação é condição necessária.

Nesse sentido, o direito à explicação deve ser compreendido em conjunto com os outros instrumentos que definem o poder de regulação da autoridade de proteção de dados previstos principalmente em seu art. 58, como o poder de determinar uma auditoria (*data protection audit*). As auditorias são mecanismos de fiscalização do tratamento de dados nos contextos que as explicações não seriam suficientes ou que a transparência se encontra prejudicada por direitos associados aos negócios ou aos fins do processamento de dados.

---

[602] RAJI, I. D. *et al.* Closing the AI accountability gap: defining an end-to-end framework for internal algorithmic auditing. *In*: FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency. New York, USA: Association for Computing Machinery, 2020. p. 33–44. Disponível em: https://doi.org/10.1145/3351095.3372873. Acesso em: 1 set. 2020.

[603] RUDIN, C. Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead. *ArXiv [cs, stat]*, [*S. l.*], 2019. Disponível em: http://arxiv.org/abs/1811.10154. Acesso em: 10 ago. 2020.

Podemos notar essa associação entre explicação, auditoria e o papel fiscalizador da autoridade nacional também na LGPD. O art. 20 da lei, que trata das decisões automatizadas, estabelece em seu *caput* o direito à revisão de decisões tomadas de forma unicamente automatizada e define em seu § 1º o direito à explicação, limitado pelos direitos associados aos segredos de negócio. A limitação para segredos de negócio poderia tornar toda a previsão de explicação e revisão inócuas, não fosse o disposto no § 2º: quando não fornecidas as informações sobre o processamento de dados sob o argumento de segredo comercial e industrial, a autoridade nacional poderá realizar uma auditoria para verificação de aspectos discriminatórios no tratamento de dados.

Além daquela previsão da LGPD para realização de auditorias para os casos de decisões automáticas para verificar a existência de discriminação, o uso desse instrumento pode ser invocado também em outros contextos. O art. 55-J da lei, em seu inciso IV, define a competência da ANPD para "fiscalizar e aplicar sanções". Em relação a esse dever de fiscalização, o inciso XVI permite a autoridade "[...] realizar auditorias, ou determinar sua realização [...] sobre o tratamento de dados pessoais efetuado pelos agentes de tratamento, incluído o poder público."[604]

O Decreto nº 10.474 de 2020, que estabelece e regulamenta a ANPD, retoma as atribuições da autoridade já estipuladas na LGPD, incluindo as previsões de auditorias, tanto para verificar a existência de discriminação nos casos em que a explicação seja negada aos titulares por conta de segredos comerciais e industriais ou como prerrogativa decorrente do poder de fiscalização da autoridade.

Segundo o regimento interno da ANPD, as auditorias são de competência da Coordenação de Geral Fiscalização, com apoios da Coordenação Geral de Pesquisa e Tecnologia, conforme os arts. 17 e 18 da Portaria nº 01, de 8 de março de 2021. O texto repete o disposto na LGPD sobre a possibilidade de realização pela própria entidade ou por terceiros. O regimento, contudo, não avança na definição de quais seriam as entidades capazes de realizar tais auditorias e nem os procedimentos e critérios precisos sobre sua dimensão e profundidade.

Ainda não é possível fornecer uma resposta definitiva sobre essa questão, e, quanto a isso, trabalhos e publicações recentes sobre as auditorias de algoritmo têm

---

[604] BRASIL. *Lei nº 13.709, de 14 de agosto de 2018*. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF: Presidência da República, 14 ago. 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015- 2018/2018/lei/L13709.htm. Acesso em: 07. nov. 2020.

destacado a imprecisão do termo e os possíveis contornos que pode assumir a partir da atuação das autoridades de proteção de dados.

O relatório do *Ada Lovelace Institute* e do *Data Kind* UK, *"*Examining the Black Box: Tools for Assessing Algorithms Systems"[605], discute o significado que o termo auditoria de algoritmo apresenta em vários trabalhos recentes sobre o tema. O relatório ainda discorre sobre as formas de avaliação de impacto e formas de desenvolvimento de tecnologias mais confiáveis. O relatório destaca como essas abordagens ainda estão em construção e que os termos são usados muitas vezes como sinônimos de outras práticas. Sobre as auditorias, o relatório destaca dois tipos principais: a) auditoria de viés e b) inspeção regulatória.

Ambas as abordagens surgem no contexto de demanda por maior *accountability*, principalmente nas grandes plataformas tecnológicas. A primeira abordagem é mais específica e analisa os *inputs* e *outputs*. A segunda é uma inspeção ampla e inclui mais elementos como finanças e outros atributos da regulação. Ambas as abordagens têm práticas que podem ser utilizadas para acessar os algoritmos como forma de controle externo como pela sociedade civil, pela mídia ou por órgãos reguladores. O relatório destaca ainda que nada impede que tais auditorias sejam realizadas internamente.

Diante de seu uso disseminado para práticas comerciais e dos riscos envolvidos para direitos individuais e coletivos, de casos emblemáticos envolvendo o uso dessa tecnologia, tornou-se comum que estudiosos realizassem auditorias independentes nos algoritmos, inspiradas nos *audit studies* da sociologia norte-americana em busca de vieses e discriminações. Esses estudos são testes experimentais em processos possivelmente discriminatórios para responder a algumas questões específicas sobre discriminação. São empregadas técnicas de pesquisa onde se dividem dois grupos com características iguais, com exceção de uma, qual seja, aquela que se pretende avaliar. Essa categoria pode ser sexo, raça, nacionalidade etc. Submetem-se esses grupos aos processos e avalia-se os resultados em busca de disparidades no tratamento. Tais estudos permitem observar desigualdades no mercado de trabalho, no acesso a serviços ou educação.

As auditorias de viés são modelos de auditoria parecidos àqueles estudos das ciências sociais. Podem ser realizados testes no sistema com entradas homogêneas

[605] ADA INSTITUTE. *Examining Tools for assessing algorithmic systems the Black Box*. [*S. l.*]: [*s. n.*], 2020.

e variações nas características étnicas, raciais, de gênero, religiosas e analisar as saídas. Em geral são realizadas por pessoas não envolvidas na construção do algoritmo e não precisam acessar o código da aplicação, avaliando apenas os resultados, de forma que pode ocorrer sem o conhecimento dos controladores do algoritmo, o que implica em questões éticas e legais. É um modelo que requer um teste de hipótese bem delimitado, com questões claras e bem formuladas, de forma que se possa utilizar técnicas básicas de estatística, porém robustas, para avaliar de maneira efetiva a isenção do modelo algorítmico.

Se no campo da sociologia as análises visam fenômenos gerais, não relacionados a uma pessoa ou corporação, no campo de um algoritmo geralmente visam um alvo específico. Por isso se assemelham mais a investigações e acusações criminais. Nesse sentido, apresentam algumas limitações e riscos. Podem afetar os direitos das empresas, causar danos e comprometer a segurança dos sistemas[606]. O futuro dessa metodologia envolve o desenvolvimento de meta-análises sobre seus efeitos na governança das empresas e sua efetividade como forma de controle externo. Além disso, é preciso pensar em novas aplicações para novos cenários de *machine learning*, o que implica o desenvolvimento de modelos de auditoria contínua.

A outra forma de auditoria é mais parecida com as auditorias clássicas dos meios corporativos. São inspeções regulatórias e operacionais. Tais empreendimentos analisam todo o ciclo do sistema, incluindo a adequação à proteção de dados pessoais, aos demais direitos civis e outras regulações setoriais. Por sua abrangência depende de entidades com poder regulatório para serem autorizadas ou de apoio interno da organização auditada. Sua implementação é sempre muito contextualizada, de forma que não é possível traçar metodologias padronizadas. Dentre os vários procedimentos possíveis para utilizar, tal inspeção pode incluir uma auditoria de viés.

Tais auditorias podem ser realizadas por entidades regulatórias ou por auditores. Pela sua dimensão e pelo acesso externo às informações sensíveis das empresas, tais inspeções precisam de um terceiro confiável, com conhecimento sobre questões de programação e sobre o contexto de aplicação da tecnologia. O que implica no desenvolvimento de um arranjo institucional e a produção de um corpo de conhecimentos pelos reguladores, que inclui criação de um conjunto de poderes que

---

606 SANDVIG, C. *et al.* Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms. [*S. l.*]: [*s. n*.], 2014. p. 23.

constituam um mandato para realizar inspeções de maneira previsível e segura para empresas, de forma a tornar possível o desenvolvimento de sistemas auditáveis[607].

As formas de auditoria devem variar de acordo com o contexto, visto que qualquer técnica fornece uma visão apenas parcial. A auditoria nos códigos, por exemplo, à primeira vista pode parecer como a mais eficiente forma de se avaliar a atuação de um algoritmo. No entanto, como aponta Sandvig *et. al.*, os algoritmos das empresas se assemelham a um verdadeiro quebra cabeça de códigos diversos, controlados por um time de engenheiros, de forma que não podem ser facilmente interpretados.

Além disso, a análise dos códigos lógicos diz pouco sobre como uma aplicação opera na prática. Muitos modelos dependem da composição de dados utilizados em seu treinamento, seus códigos consistem em meras rotinas matemáticas como regressões. Nesse sentido, uma análise da representatividade dos dados é essencial para entendermos a adequação do algoritmo. Da mesma forma, a análise dos códigos e os dados não permite avaliar as implicações sobre direitos desde a coleta dos dados, ou mesmo sobre as finalidades do tratamento de dados. Tais análises também não permitem tirar conclusões sobre a robustez do sistema e sobre aspectos organizacionais e processos internos do controlador da tecnologia para preservar a segurança dos dados, os direitos dos titulares e a capacidade de o sistema responder a imprevistos, ataques ou incidentes de segurança.

A emergência de novas aplicações e a maior complexidade de sistemas de decisão automática já movimentam os reguladores no sentido de desenvolverem arranjos e estruturas de auditorias, principalmente relacionadas ao contexto de algoritmos de inteligência artificial. O *Guidance on the AI auditing framework: Draft guidance for consultation*[608]*,* documento preparatório e consultivo do ICO., apresenta algumas ferramentas de auditoria e de adequação à proteção de dados desde o desenvolvimento até a implementação de sistemas de inteligência artificial. Após a consulta pública sobre o tema, a autoridade pretende consolidar um corpo de boas práticas destinadas aos desenvolvedores ou àqueles que utilizam tecnologias de terceiros. Tais práticas deverão guiar as futuras auditorias da entidade.

---

[607] ADA INSTITUTE. *Examining Tools for assessing algorithmic systems the Black Box*. [*S. l.*]: [*s. n.*], 2020.

[608] INFORMATION COMMISSIONER'S OFFICE. *Guidance on the AI auditing framework*. [*S. l.*]: ICO, 2020. Disponível em: https://ico.org.uk/media/about-the-ico/consultations/2617219/guidance-on-the-ai-auditing-framework-draft-for-consultation.pdf. Acesso em: 1 ago. 2020.

O guia divide-se em três etapas para uma análise da inteligência artificial. A primeira envolve uma análise de *accountability*, o que requer a adequação à legislação de proteção de dados, a existência de documentos como análises de impacto, cumprimento adequado de obrigações de operador e controlador e o correto sopesamento de princípios de proteção de dados para a aplicação como um todo e da criação de mecanismos eficientes de governança. Nessa etapa trata-se de avaliar a conhecimento, distribuição e responsabilização pelos riscos dos sistemas, que não podem ser meramente delegados para equipes técnicas.

Nessa avaliação é preciso mapear os fluxos e processos, a diversidade das equipes e o treinamento adequado em toda a cadeia, incluindo os parceiros e relações contratuais. O documento adota uma abordagem baseada em riscos, de forma a avaliar se os instrumentos de governança são proporcionais aos riscos envolvidos no uso da inteligência artificial e se são adequadamente enfrentados em termos de mitigação e de sua proporcionalidade aos interesses da organização e dos titulares dos dados.

A segunda parte do documento dedica-se a apresentar medidas para avaliar a justiça, legalidade e a transparência do processamento. As medidas propostas pelo guia envolvem a documentação de todas as decisões tomadas e alguns cuidados no desenvolvimento e implementação, como a definição de uma aplicação e finalidades específicas, como forma de fornecer uma base robusta para os cálculos de risco. Essas decisões devem estar embasadas na escolha correta das bases legais dos regulamentos de proteção de dados.

A terceira parte dedica-se à avaliação dos princípios de segurança e minimização de dados nos contextos da inteligência artificial. A minimização é essencial para evitar uma série de riscos envolvendo tais aplicações. Alguns desses riscos representam novidades nos meios tecnológicos. Alguns deles envolvendo usos maliciosos que tem um potencial de impacto muito alto em decorrência da escala.

Algumas reflexões já surgem no sentido de desenvolverem-se metodologias de auditoria interna para as organizações que desenvolvem ou utilizam algoritmos[609]. Há um argumento importante no sentido de que tais práticas de controle de processos

[609] RAJI, I. D. *et al.* Closing the AI accountability gap: defining an end-to-end framework for internal algorithmic auditing. *In*: FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency. New York, USA: Association for Computing Machinery, 2020. p. 33–44. Disponível em: https://doi.org/10.1145/3351095.3372873. Acesso em: 1 set. 2020.

podem ajudar a aumentar a confiança nas tecnologias e a credibilidade das organizações. No entanto, não se pode deixar de relacionar que tais reflexões surgiram apenas após um período de questionamentos na academia e na opinião pública sobre os riscos no uso de algoritmos e de o GDPR consolidar maior poder para as autoridades de proteção de dados.

5.2 ORIENTAÇÕES PRÁTICAS PARA A GARANTIA DO DIREITO À EXPLICAÇÃO: PROPOSTA DE UM *FRAMEWORK* DE EXPLICABILIDADE A PARTIR DA CLÁUSULA GERAL DO DEVIDO PROCESSO INFORMACIONAL

No presente capítulo, buscaremos fornecer uma abordagem prática para a implementação do direito à explicação no contexto do devido processo informacional, a partir da articulação e sistematização de elementos técnico-jurídicos apresentados nos capítulos anteriores. Antes de apresentar o *framework* de explicabilidade propriamente dito, na seção 5.2.1 apresentamos algumas diretrizes e pressupostos balizadores do arcabouço proposto.

Assim, na subseção 5.2.1.1, partindo-se das considerações elaboradas no Capítulo 2, em que se buscou demonstrar que há diferentes modalidades de explicação, inseridas em contextos distintos e direcionadas a diferentes objetivos e tipos de sujeitos, pretende-se argumentar que a implementação do direito à explicação não depende de uma fórmula pronta ou genérica (*one size fits all*). Nesse sentido, buscaremos demonstrar que há uma “caixa de ferramentas”, composta por diferentes instrumentos técnico-jurídicos, apresentados no tópico 5.1, que podem ser empregados para a implementação do direito à explicação. Cada uma destas ferramentas carrega potencialidades e limitações, sendo mais ou menos adequadas em função dos objetivos que se busca atingir, do contexto considerado e dos destinatários da explicação, não havendo uma receita pronta, mas um amplo ferramental aplicável a uma infinidade de casos.

Na subseção 5.2.1.2, desenvolvemos o argumento de que um modelo de explicabilidade deve ser concebido levando em conta a opacidade e os riscos de cada atividade, e de que uma avaliação de risco deve estar contextualmente situada, pois há diferentes riscos para cada domínio de aplicação considerado. Nesse sentido, discutimos, a partir da bibliografia e da legislação vigente, o conceito de risco associado a domínios e a aplicações específicas e o conceito de opacidade inerente aos modelos de decisões automatizadas, de forma a orientar a instrumentalização da

mitigação de riscos, da explicação e do devido processo informacional. Para ilustrar a operacionalização dessas variáveis, trabalharemos com um diagrama de análise de risco e impacto que considere sistemas com diferentes níveis de complexidade e opacidade em diferentes domínios. Para tanto, serão considerados 3 diferentes sistemas em três domínios específicos: (i) segurança pública e persecução criminal, (ii) relações de emprego e prestação de serviços mediadas por plataformas e (iii) concessão de benefícios sociais.

Na subseção 5.2.1.3, desenvolvemos o argumento de que um modelo de explicação deve ser concebido em função do destinatário da explicação. Para tanto, propomos que, ao se analisar o destinatário da explicação, quatro dimensões devem ser consideradas: (i) contexto da interação, (ii) necessidades, (iii) complexidade do sistema e (iv) possibilidade de desafiar a decisão.

Por fim, na subseção 5.2.2, articulamos as considerações preliminares apresentadas na subseção 5.2.1 e propomos um *framework* de explicabilidade como referência para a construção de modelos de explicabilidade que considerem a cláusula geral do devido processo informacional, os riscos de cada domínio de atividade, a opacidade dos sistemas e os destinatários da explicação, em suas quatro dimensões, conforme apresentadas no item 5.2.1.3. Para tanto, primeiramente desenvolvemos as linhas gerais do que chamados de cláusula geral do devido processo informacional, composta por cinco qualificadoras — isenção, informação, compreensão, recorribilidade e revisão — somadas ao princípio da prevenção, entendido como seu eixo transversal e estruturante, e buscamos contextualizá-las traçando um paralelo com a cláusula geral do devido processo legal. Após a apresentação das cinco qualificadoras, nos atemos a duas delas, informação e compreensão, por entendermos que estão mais diretamente relacionados à construção de uma explicação, e de como elas podem ser colocadas em movimento a partir da ideia de risco, opacidade, explicação centrada no destinatário e da caixa de ferramentas técnico-jurídicas existente.

5.2.1 Diretrizes e pressupostos balizadores do *framework* de explicabilidade proposto

5.2.1.1 Uma caixa de ferramentas para a garantia do direito à explicação

A primeira consideração a ser levada em conta para a compreensão do

framework de explicabilidade aqui proposto é a de que ele não consiste numa fórmula pronta ou genérica (*one size fits all)*, aplicável diretamente à generalidade de casos. Como veremos, cada explicação demandará a articulação de instrumentos técnico-jurídicos distintos, a depender do perfil e das necessidades de seus destinatários, do domínio de aplicação do sistema automatizado em questão, do grau de risco envolvido e das expectativas em torno da explicação a ser fornecida. Para além da proposição de um modelo fixo e generalista, faz-se necessário conceber um modelo flexível, adaptável às especificidades e necessidades subjacentes a cada caso concreto.

Neste sentido, o *framework* de explicabilidade proposto parte da constatação de que não há uma solução única, mas sim uma "caixa de ferramentas", composta por diferentes instrumentos técnico-jurídicos, apresentados no tópico 5.1, que podem ser combinados e empregados para a construção de modelo de explicação em uma determinada situação concreta. Ademais, é preciso considerar que cada um dos instrumentos à disposição na caixa de ferramentas possui potencialidades e limitações que lhes são inerentes, o que os torna mais ou menos adequados em função dos objetivos que se pretende atingir, do contexto considerado e dos destinatários da explicação. Deve-se partir do pressuposto, portanto, de que não há uma receita pronta, mas sim um amplo ferramental aplicável a uma infinidade de casos.

5.2.1.2 Por uma abordagem contextual: risco e opacidade do sistema como variáveis norteadoras de um modelo de explicabilidade

Nesta seção propomos um quadro que deve auxiliar nas avaliações de risco de uma organização, no contexto de aplicação e visando a implementação do direito à explicação. Nossa proposta enquadra as aplicações em dois eixos centrais para a análise contextual e para que a organização possa desenvolver as medidas necessárias para garantir o direito à explicação. O primeiro eixo consiste no grau de risco de uma aplicação, o segundo consiste no grau de opacidade do sistema. Cabe destacar que esse quadro não substitui outras avaliações e metodologias de avaliação de riscos, servindo apenas como um subsídio para o desenvolvimento de um framework de explicabilidade, subsidiariamente ou de forma complementar aos demais processos de proteção de dados de uma organização.

No item 5.1.3 tratamos sobre o princípio da *accountability* e no item 5.1.4 dos

relatórios de impacto. Nestas duas oportunidades apresentamos uma discussão mais aprofundada sobre o conceito de risco e sua importância para a regulação da proteção de dados. O conceito que apresentaremos aqui tem um propósito instrumental, como forma de operacionalização de uma avaliação inicial que possa orientar o desenvolvimento da explicabilidade. As avaliações de riscos presentes nos diversos tipos de relatórios de impacto devem ser muito mais detalhadas e exaustivas. A realização desses relatórios de impacto, inclusive, como argumentamos na próxima seção, pode representar uma tarefa essencial para o desenvolvimento de um framework de explicabilidade.

A definição de risco adotada pelo GDPR corresponde à possibilidade de ocorrência de um evento indesejado[610]. Algumas organizações, incluindo autoridades nacionais de proteção de dados, apontam *templates* e metodologias para avaliação de riscos. Discutimos na seção 5.1.4, que trata sobre os relatórios de impacto, como a regulação europeia baseia-se no conceito de probabilidade e de severidade do evento indesejado para classificar o grau de risco de um processamento de dados como alto[611]. Nessa escala, como vimos, não há uma classificação numérica, mas uma classificação categórica em função da junção dessas duas variáveis[612].

E é uma classificação como essa que podemos adotar para nossa análise. Nos baseamos nos critérios e exemplos apresentados na regulação, discutidos pela literatura e propostos por autoridades nacionais[613] para construção de um quadro que

---

[610] GELLERT, R. Understanding data protection as risk regulation. *Journal of Internet Law*, [*S. l.*], p. 3–15, 2015.

[611] GOMES, M. C. Relatório de impacto à proteção de dados. Uma breve análise da sua definição e papel na LGPD. *Revista do Advogado*, [*S. l.*], v. 39, n. 144, p. 174–183, 2019; e DEMETZOU, K. Data Protection Impact Assessment: A tool for accountability and the unclarified concept of 'high risk' in the General Data Protection Regulation. *Computer Law & Security Review*, [*S. l.*], v. 35, n. 6, p. 105342, 2019. Disponível em: https://doi.org/10.1016/j.clsr.2019.105342. Acesso em: 21 jun. 2021.

[612] A classificação de alto risco proposta pelo GDPR está vinculada à obrigação de realização do DPIA prevista no art. 35. O dispositivo prevê a obrigatoriedade da realização do relatório para aplicações de alto risco. Neste sentido, o regulamento define, corretamente, o conceito de alto risco de uma forma bem ampla, para que essa obrigação seja mandatória. No entanto, nos próprios exemplos traçados nos recitais é possível depreender graus diferenciados de risco.

[613] EDPS. *Survey on Data Protection Impact Assessments under Article 39 of the Regulation*. [*S. l.*]: EDPS, 2020. Disponível em: https://edps.europa.eu/sites/edp/files/publication/20-07-06_edps_dpias_survey_en.pdf. Acesso em: 9 jun. 2020; CNIL. *Privacy Impact Assessment Methodology*. [*S. l.*]: CNIL, 2018. Disponível em: https://www.cnil.fr/sites/default/files/atoms/files/cnil-pia-1-en-methodology.pdf. Acesso em: 9 jun. 2020; OAIC. *Privacy Impact Assessment Guide*. Sydney: OAIC, 2010. Disponível em: http://www.icb.org.au/out/?dlid=38156. Acesso em: 14 abr. 2021; AGENCIA DE ACCESO A LA INFORMACIÓN PÚBLICA; UNIDAD REGULADORA Y DE CONTROL DE DATOS PERSONALES. *Guía de Evaluación de Impacto en la Protección de Datos*. [*S. l.*]: Agencia de acceso a la información pública; Unidad reguladora y de control de Datos Personales, 2020. Disponível em: https://www.argentina.gob.ar/sites/default/files/guia_final.pdf. Acesso em: 9 jun. 2021; e

apresenta três níveis de riscos aos direitos dos titulares: baixo, moderado e alto[614]. As aplicações de baixo risco consistem nos modelos cujos efeitos indesejados são improváveis e muito baixos, ou seja, não causam mais que um mero aborrecimento e podem ser revertidos facilmente. As aplicações de nível moderado de risco consistem naquelas cujas probabilidades de eventos indesejados seja factível ou cujos efeitos sejam relevantes, afetem de forma moderada a propriedade, o acesso a bens, direitos ou a reputação, mas que possam ser facilmente revertidos. Utilizaremos, por fim, a categoria de alto risco, nas quais há grande probabilidade de falhas, ou essas falhas possam afetar os sujeitos ou grupos de maneira muito severa em seus direitos fundamentais, liberdades, saúde física e mental, de maneira que as exigências para o devido processo são mais prementes, como nos campos da segurança pública, da saúde, dos direitos sociais ou mesmo em outras aplicações com alto potencial de gerar efeitos sistêmicos com potenciais discriminatórios.

Discutimos no Capítulo 4 diferentes tipos de opacidade, baseados na classificação proposta por Jenna Burrell que divide a opacidade em 3 categorias diferentes[615]. Uma delas derivada das necessidades comerciais. As outras duas, que nos interessam mais nessa seção, derivam de características técnicas ou cognitivas. De um lado, há a opacidade derivada da natureza técnica dos algoritmos como um campo específico do conhecimento, qual seja, a programação e tecnologia da informação. A outra derivada da própria natureza não interpretável de alguns resultados de aplicações de inteligência artificial. Discutimos essas questões na seção 4.2.1

Neste sentido, podemos conceber ao menos 3 níveis de opacidade. O primeiro deles seria a ausência de opacidade, em aplicações facilmente compreensíveis, cujas entradas, os processos intermediários e os resultados possam ser conhecidos de forma imediata. O segundo nível de opacidade decorre de aplicações mais complexas cujos dados de entrada, os processos intermediários ou os resultados demandam conhecimentos técnicos específicos sobre áreas do conhecimento para serem

---

AEPD. Guía práctica para las evaluaciones de Impacto en la Protección de los datos sujetas al RGPD. *AEPD*, 6 maio 2021. Disponível em: https://www.aepd.es/es/derechos-y-deberes/cumple-tus-deberes/medidas-de-cumplimiento/evaluaciones-de-impacto. Acesso em: 6 jun. 2021.

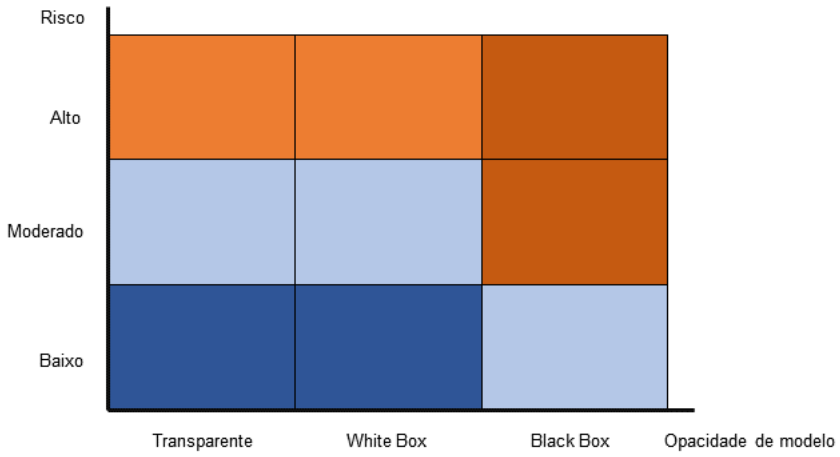
614 Um documento da *Unidad Reguladora y de Control de los Datos Personales*, a autoridade de proteção de dados da Argentina, oferece um modelo de matriz de riscos nos quais é possível encontrar metodologias de classificação para níveis de probabilidade e de impacto.

615 BURRELL, J. *How the Machine "Thinks:" Understanding Opacity in Machine Learning Algorithms*. Rochester, NY: Social Science Research Network, 2015. Disponível em: https://papers.ssrn.com/abstract=2660674. Acesso em: 27 maio. 2020.

compreendidos. As aplicações com essas características podem ser compreendidas como modelos *white box*.

O terceiro nível de opacidade é típico de modelos conhecidos como *black box*. São aplicações não interpretáveis, das quais só podemos conhecer os dados de entrada e os dados de saída, mas nas quais não é possível conhecer os processos intermediários e os critérios usados para se atingir os resultados, como são alguns modelos de *machine learning*. Essas diferentes escalas de opacidade demandam diferentes medidas para a explicabilidade.

GRÁFICO 8 – Escala de risco e opacidade



Fonte: Elaboração própria

Apresentaremos a seguir a aplicação desse quadro a partir de 3 casos de diferentes aplicações em diferentes domínios., de forma a ilustrar como é possível compreender um sistema de decisão automatizada segundo os eixos de risco e opacidade. Apresentaremos 3 aplicações: a) o uso de reconhecimento facial para fins de persecução criminal da área de segurança pública; b) o uso de algoritmos em plataformas de trabalho ou prestação de serviços; e o c) uso de algoritmos na análise de fraudes na concessão de benefícios sociais.

A) Reconhecimento facial na área da segurança pública

As tecnologias de reconhecimento facial se popularizaram nos últimos anos. A tecnologia, modalidade do que se convenciona chamar de visão computacional, encontra diversos usos nos meios privados e públicos. A visão computacional é um componente comum no desenvolvimento de veículos autônomos, por exemplo. Além disso, muitos estados têm adotado a tecnologia em seus programas de segurança pública. O tema tem levantado debates em entidades da sociedade civil, autoridades públicas e empresas.

As aplicações de reconhecimento facial operam a partir da utilização de câmeras em locais de utilização públicos. Podem ser alocadas nas ruas, nas estações de trem ou metrô, nos terminais de ônibus etc. Essas câmeras captam imagens dos indivíduos e comparam tais imagens com um banco de dados, composto por outras imagens previamente coletadas. A partir disso, um algoritmo determina se um indivíduo corresponde a uma pessoa investigada ou condenada por alguma infração penal.

Entre os riscos apontados para o uso da tecnologia, há a possibilidade de que falsos positivos submetam pessoas inocentes a inquéritos e processos criminais. As tecnologias são testadas e possuem um grau de acurácia e uma porcentagem de erros reconhecidas. No entanto, ao se aplicar em massa, é certo que o número de erros será significativo, de forma que seja necessário utilizar tais aplicações com cautela. Dito de outra forma, a escala de uso de tais aplicações resulta em uma probabilidade alta de violação de direitos. Nesse sentido, é prudente considerar que o uso dessas tecnologias apresenta uma probabilidade alta de ocorrência de um evento indesejado.

Esse problema se agrava visto que tais aplicações de reconhecimento de imagens possuem diferentes taxas de erros para diferentes grupos sociais, principalmente em função de características raciais. Há diversos estudos que apontam taxas de erros significativamente diferentes entre pessoas brancas e negras[616].

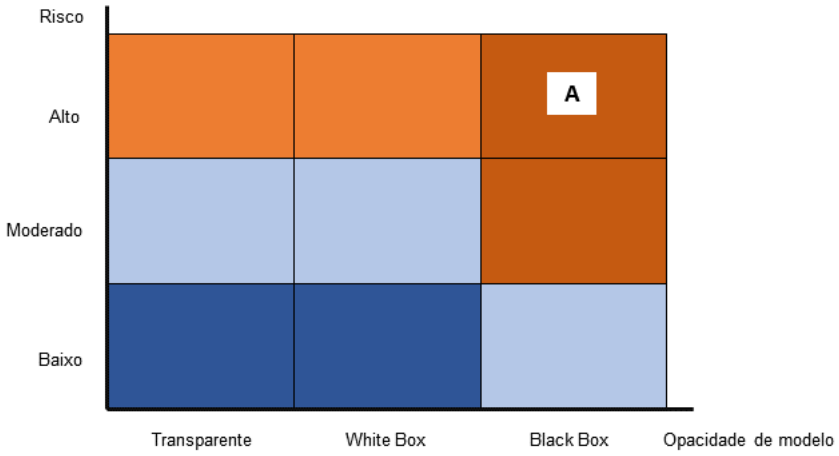
Mesmo que uma pessoa falsamente considerada suspeita não seja condenada mediante o uso de outras provas que a inocentem, a mera submissão a procedimentos

[616] PERKOWITZ, S. The Bias in the Machine: Facial Recognition Technology and Racial Disparities. *MIT Case Studies in Social and Ethical Responsibilities of Computing*, [*S. l.*], 2021. Disponível em: https://doi.org/10.21428/2c646de5.62272586. Acesso em: 29 maio 2021; e LESLIE, D. Understanding bias in facial recognition technologies. *ArXiv [cs]*, [*S. l.*], 2020. Disponível em: https://doi.org/10.5281/zenodo.4050457. Acesso em: 29 maio 2021.

de investigação criminal representa um impacto relevante. Nesse sentido, pela esfera das liberdades individuais afetadas, e pelo potencial discriminatório em desfavor da população negra, devemos considerar que a severidade do evento indesejado seja alta.

Do ponto de vista da transparência, quando analisamos a tecnologia em que se baseiam as aplicações de reconhecimento facial, encontramos um grau muito alto de opacidade. A metodologia por trás do reconhecimento facial é denominada rede neural, um modelo não supervisionado de aprendizado de máquina, no qual é possível conhecermos os dados de entrada e os resultados (*outputs*) do sistema, mas cujos processos intermediários realizados pela máquina não são compreensíveis ou exprimíveis de forma lógica. Em outras palavras, o sistema tem um resultado no qual afirma que determinada imagem de indivíduo tem uma probabilidade de corresponder a uma pessoa procurada ou suspeita. Embora possamos saber o grau de acurácia desse resultado, não é possível compreender, a priori, como o algoritmo chegou a tal resultado de forma causal.

GRÁFICO 9 – Escala de risco e opacidade II



Fonte: Elaboração própria

Quando analisamos essas duas dimensões de forma conjunta, podemos ter uma classificação útil para pensarmos na mitigação de riscos e na explicabilidade.

Aplicações opacas por razões técnicas e com impactos significativos demandam esforços relevantes das organizações para garantir os direitos dos titulares e a transparência. A etapa de desenvolvimento precisa basear-se em estudos robustos e representativos dos resultados, e, além disso, é preciso fornecer informações relevantes para os sujeitos afetados por tais decisões. É importante considerar, outrossim, os efeitos que tais aplicações possuem para grupos sociais específicos, e não apenas para indivíduos determinados, de forma que a explicabilidade deve levar em consideração os interesses coletivos, que também devem ser tutelados pela proteção de dados[617].

No campo da segurança pública e da justiça criminal, as decisões automatizadas devem levantar um sinal claro de alerta para altíssimos riscos, principalmente frente a aplicações protegidas por segredos comerciais. Quando nos deparamos com algoritmos não interpretáveis como os utilizados na visão computacional, esse alerta deve ser ainda reforçado. Organizações da sociedade civil e estudiosos do tema defendem o banimento do uso do reconhecimento facial para qualquer tipo de vigilância em massa[618]. Na persecução criminal, há um conflito entre a natureza do resultado do algoritmo e o devido processo, na utilização de provas das quais não se pode conhecer os critérios que fundamentam a conclusão. Algumas empresas de tecnologia interromperam o fornecimento e o desenvolvimento dessas tecnologias em face de tais riscos.

No caso brasileiro, algumas cidades têm utilizado o reconhecimento facial e já há casos de pessoas erroneamente detidas em função de seu uso[619]. Caso se compreenda que o uso de reconhecimento facial seja necessário e possível na segurança pública, alguns apontamentos são importantes. Ainda que a LGPD, em seu Art. 4º, II, remeta que a segurança pública será regida por outra legislação, o §1° do mesmo artigo ressalta que o processamento de dados pessoais para fins de

---

[617] ZANATTA, R. A. F. A tutela coletiva na proteção de dados pessoais. *Revista do Advogado*, n. 144, nov. 2019. Disponível em: https://aplicacao.aasp.org.br/aasp/servicos/revista_advogado/paginaveis/144/200/index.html. Acesso em: 20 maio 2021. p. 201-208; ZANATTA, R. A. F. Tutela coletiva e coletivização da proteção de dados pessoais. *In:* PALHARES, F. (Org.). *Temas Atuais de Proteção de Dados*. São Paulo: Thomson Reuters, 2020. p. 345–374.

[618] DIGITAL rights alliance file legal complaints across Europe against facial recognition Company Clearview AI. *Noyb*, 26 maio 2021. Disponível em: https://noyb.eu/en/digital-rights-alliance-file-legal-complaints-against-facial-recognition-company-clearview-ai. Acesso em: 29 maio 2021.

[619] INSTITUTO IGARAPÉ. *Reconhecimento Facial no Brasil*. INFOGRÁFICO RECONHECIMENTO FACIAL NO BRASIL — INSTITUTO IGARAPÉ. Instituto Igarapé, 2020. Disponível em: https://igarape.org.br/infografico-reconhecimento-facial-no-brasil/. Acesso em: 29 maio 2021.

segurança pública deve observar os princípios da proteção de dados e os direitos dos titulares previstos na lei. Nesse sentido, estudos adicionais sobre o uso dessa tecnologia neste setor seriam de extremo valor para a disciplina da proteção de dados e dos direitos fundamentais.

B) Plataformas de fornecimento e prestação de serviços de entrega

A popularização de dispositivos equipados com tecnologias de geolocalização e conexão à internet possibilitou o desenvolvimento de uma série de plataformas que conectam usuários e prestadores de serviço, como motoristas, estabelecimentos comerciais e entregadores. Uma série de trabalhos têm discutido os impactos dos modelos de negócios criados por essas plataformas, em termos jurídicos e econômicos[620].

Essas plataformas, que criaram novas formas de consumo e trabalho, utilizam-se de dados pessoais para o funcionamento e realizam uma série de operações de forma automatizada, por meio de algoritmos. A distribuição de pedidos ou de viagens é realizada de forma automatizada, baseada em informações de geolocalização. A definição do preço pago pelos clientes pelo serviço de entrega ou pela viagem, bem como o valor pago aos motoristas e entregadores, também é realizada de forma automatizada. Algumas plataformas utilizam procedimentos automatizados nos processos de bloqueio ou banimento de motoristas ou entregadores por desrespeitarem as políticas de uso da plataforma.

Sob o ponto de vista da análise de riscos, podemos identificar uma série de eventos indesejados. Em geral, o risco envolve a possibilidade de que os resultados dos algoritmos — a distribuição de pedidos, o cálculo dos valores, a constatação de má conduta do motorista ou entregador — não correspondam às disposições contratuais ou às legislações civis ou trabalhistas. Esses erros podem derivar de inconsistências materiais relativas aos dados utilizados na plataforma, ou mesmo aos critérios utilizados na programação dos sistemas, que podem implicar em conflitos de interesse entre entregadores e plataformas.

A falta de equidade na distribuição das corridas ou o bloqueio indevido de motoristas e entregadores tem o potencial de afetar a renda e as condições de vida

[620] ZANATTA, R. A. F.; PAULA, P. C. B. de; KIRA, B. *Economias do Compartilhamento e o Direito*. 1. ed. São Paulo: Juruá Editora, 2017.

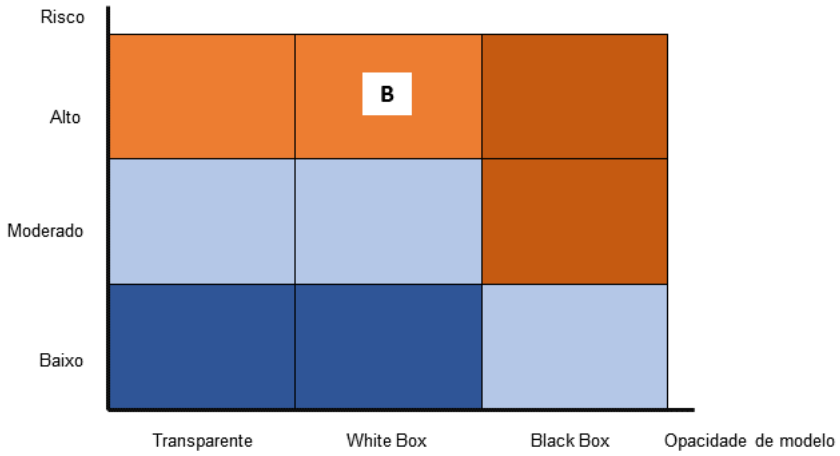
dos prestadores de serviço ou trabalhadores. O uso de critérios que impliquem em diferenciação, privilégio ou benefícios diferenciais pode ter efeitos sistêmicos relevantes, como escamotear relação de subordinação, próprias das relações de emprego, alijando os usuários de direitos relevantes. Além disso, o fornecimento de incentivos e estímulos a partir de critérios de produtividade pode incentivar motoristas e entregadores a realizarem jornadas excessivas de trabalho, colocando-os em situações de risco.

Por afetar diretamente as condições de vida e trabalho dos sujeitos, podemos considerar esses riscos relevantes do ponto de vista da regulação da proteção de dados, de forma que podemos classificar os riscos como altos. No entanto, cabe destacar que as avaliações de risco não são afirmações perenes. É possível que eventos e efeitos sistêmicos desconhecidos justifiquem uma classificação mais alta.

Os algoritmos utilizados por essas plataformas são passíveis de interpretação, como podemos apreender das próprias especificações e comunicados das empresas. São critérios de distância, valores monetários, tempo de conexão, avaliação dos usuários, tempo de corrida, valor dos pedidos, valor da corrida, taxas de desconto ou promoção. A partir desses dados são modulados resultados possíveis buscando a maior eficiência econômica para as plataformas. Nesse sentido, o conhecimento do código fonte de tais aplicações pode fornecer as informações necessárias para sua compreensão. São algoritmos *white box*.

No entanto, pela natureza dos dados e pela complexidade do processamento, há um grau de opacidade decorrente da natureza técnica da aplicação, que exige conhecimentos de programação, matemática, de computação etc. Há variáveis e índices utilizados que demandam conhecimentos mais robustos, de forma que podemos classificar o sistema no nível intermediário de opacidade.

GRÁFICO 10 – Escala de risco e opacidade III



Fonte: Elaboração própria

Cabe apontar, contudo, que tais aplicações apresentam na prática, um grau de opacidade mais elevado devido ao segredo comercial. Discutimos, na seção 4.3.3, os processos que essas empresas enfrentam nos tribunais sobre o acesso a seus códigos fonte e de como os critérios utilizados nos algoritmos são matérias relevantes para a definição de relações jurídicas na esfera do direito do trabalho. Dessa forma, é importante destacar qual é o nível de opacidade permitido, em função do segredo industrial ou comercial, condição prevista no art. 20 da LGPD, em face dos direitos dos titulares potencialmente afetados.

Se a automatização de processos pode trazer ganhos de produtividade em setores econômicos, é importante levar em consideração que tais ganhos não devem ocorrer em prejuízos a garantias legais e relações jurídicas já estabelecidas nos campos do direito, como a seara trabalhista ou mesmo aspectos do direito civil. Além disso, podemos apontar que a explicabilidade pode contribuir na construção de um modelo de negócios mais confiáveis e relações mais harmoniosas entre os diversos agentes dessas cadeias econômicas.

C) Análise de fraudes na concessão de benefícios sociais

O uso de algoritmos é comum no campo das políticas públicas. No Capítulo 3 deste trabalho, mencionamos como o surgimento das regulações de proteção de dados nos anos 70 relacionavam-se à coleta de dados pelo estado no contexto da assistência social. Desde então, é seguro supor que o fluxo e volume de dados só aumentou.

A popularização de métodos e técnicas de análise preditiva por meio de algoritmos tornou-se uma tentação comum para as agências e órgãos governamentais num ambiente de pressão por redução de custos e aumento de eficiência. Governos têm utilizado algoritmos preditivos nas mais diversas áreas, para prever abusos infantis, para conceder benefícios ou para analisar fraudes em programas sociais etc.

Na nossa análise de risco e de opacidade, tomemos o exemplo das análises de fraude em sistemas de benefícios sociais. Há diferentes modelos para análise antifraude. No geral, consistem em modelos preditivos. Esses modelos podem ser aplicados em diversos setores das políticas públicas, como forma de agilizar processos.

Algumas preocupações são comuns para todos esses tipos de aplicações. A utilização de técnicas preditivas para políticas de bem-estar infantil nos EUA, por exemplo, apresenta performances preocupantes, e a maioria dos modelos analisados em um estudo recente não levam em consideração disparidades raciais e sociais[621].

Outra questão que emerge no uso de dados governamentais para análises preditivas diz respeito à privacidade contextual. Apresentamos nesse trabalho, na seção 3.1.2.3, uma disputa na corte holandesa, onde se discutia as decisões de um sistema de análise de fraude no recebimento de benefícios sociais coletava dados de forma desproporcional e violava os direitos de privacidade e proteção dos indivíduos, que não tinham conhecimento de que seus dados eram utilizados para tal finalidade.

Uma questão comum, não apenas nos modelos preditivos, mas no uso de decisões automatizadas pelo poder público em geral, diz respeito aos danos decorrentes de erros e inconsistências das bases de dados governamentais. A concessão do auxílio emergencial no contexto brasileiro também foi realizada por

[621] SAXENA, D. *et al.* A Human-Centered Review of Algorithms used within the U.S. Child Welfare System. *In*: CHI '2020, Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. New York, NY, USA: Association for Computing Machinery, 2020. p. 1–15. Disponível em: https://doi.org/10.1145/3313831.3376229. Acesso em: 26 maio 2021.

processos automáticos. Apresentamos na seção 2.2.2 os questionamentos quanto à qualidade das decisões do sistema e do potencial excludente da tecnologia, ao utilizar diversas bases de dados sem fornecer aos usuários, que tiveram o acesso negado, informações sobre as bases de dados utilizadas. Dessa forma, os sujeitos encontravam-se sem nenhuma informação sobre quais seriam os critérios determinantes para a negativa e, consequentemente, alijados da possibilidade de recorrer.

Assim como outras aplicações de decisões automatizadas, pequenos erros em escala podem potencializar os impactos para dimensões catastróficas. Um algoritmo anti fraude foi empregado pelo Estado de Michigan para analisar fraudes nos benefícios aos desempregados chamado *Michigan Integrated Data Automated System* (MiDAS). A detecção de fraudes cancelou a participação dos beneficiários dos programas, sem nenhuma supervisão humana. Durante os anos de 2013 e 2015, de todos as fraudes detectadas, uma revisão promovida pelo próprio estado detectou que as predições estavam erradas em 93% dos casos[622].

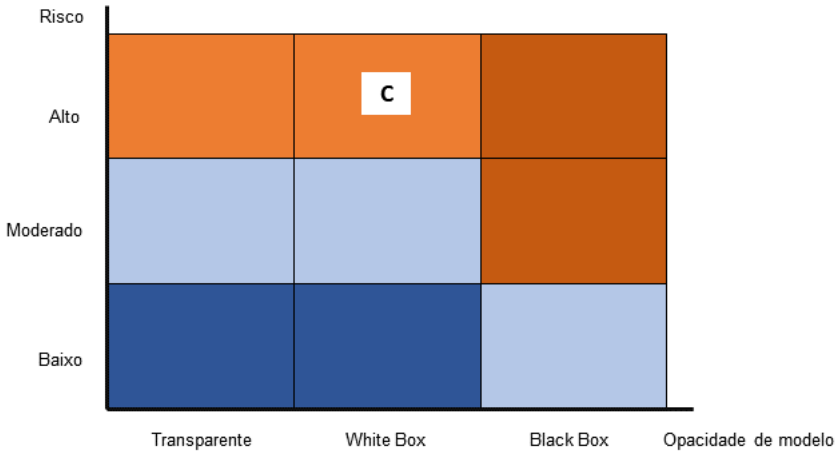
Neste sentido, do ponto de vista da probabilidade de erros, podemos concluir que é certo que tais sistemas apresentarão erros em seu funcionamento. Isso não torna tais algoritmos especiais em relação aos demais, visto que, como já é sabido, as tecnologias baseadas em estatísticas apresentam um grau de acurácia que jamais correspondem a 100%. No entanto, no campo das políticas públicas, há um risco sistêmico.

Em função de impactos diferenciais em grupos específicos, tais aplicações podem ter como efeitos a contribuição para manutenção ou agravamento de desigualdades sociais e condições estruturais de segregação. Nas situações em que correções de dados incorretos das bases de dados governamentais, ou de critérios incorretos utilizados pelo algoritmo representem custos financeiros, ou mesmo esforços muito severos, é bem possível que as populações em maior vulnerabilidade sejam duramente prejudicadas. E, nesse sentido, tais aplicações podem contribuir para manter determinados grupos sociais nas categorias estatísticas que deveriam prever, atuando como profecias auto-realizáveis. Nesse sentido, do ponto de vista da severidade dos eventos indesejáveis de tais aplicações no campo da concessão de

[622] FELTON, R. Michigan unemployment agency made 20,000 false fraud accusations – report. *The Guardian*, Detroit, 18 dez. 2016. Disponível em: http://www.theguardian.com/us-news/2016/dec/18/michigan-unemployment-agency-fraud-accusations. Acesso em: 26 maio 2021.

benefícios sociais, não há exageros em classificar tais possíveis impactos como graves.

GRÁFICO 11 – Escala de risco e opacidade IV



Fonte: Elaboração própria

Há diversos modelos que podem ser utilizados na análise de fraudes, desde modelos semi supervisionados com algum grau de opacidade técnica, o qual demandaria esforço interpretativo, ou modelos de processamento por critérios logicamente programados, cujos critérios da decisão sejam conscientemente inseridos no código. Dessa forma, podemos classificar tais aplicações com o grau de opacidade intermediário, decorrente da necessidade de conhecimentos técnicos, mas interpretável, um modelo *White Box*.

Cabe apontar que, aqui, como nos casos das plataformas de fornecimento de serviços, observamos também uma opacidade decorrente de segredos de negócio, visto que muitas delas são desenvolvidas por empresas privadas.

Não há um argumento que justifique um *trade off* entre eficiência administrativa e o devido processo legal. A implementação da explicabilidade e de mecanismos de revisão de decisões automatizadas nos campos das políticas públicas demandam atenção especial pelos possíveis impactos e pelo interesse público envolvido nos

objetivos das políticas sociais. Nesse sentido, o desenvolvimento dessas aplicações deve trazer desde o início a preocupação de estabelecer critérios para o devido processo informacional.

5.2.1.3 Instrumentalização do direito à explicação a partir de uma abordagem centrada no destinatário da explicação

Nesta seção discutiremos a instrumentalização do direito à explicação a partir dos conceitos desta tese, apresentando, a partir de uma análise concreta, como as organizações podem traçar as suas estratégias para implementação do direito. Há diversas formas de compreender a explicação. A abordagem proposta, construída a partir da literatura recente, será centrada no sujeito de direitos, em relação ao qual os instrumentos e medidas devem ser desenvolvidos.

No item 5.1.5 discutimos como se pode compreender a explicabilidade em relação ao conceito de *privacy by design* previsto no GDPR e na LGPD. Argumentamos como é importante levar em consideração a explicabilidade dos sistemas desde o início do seu desenvolvimento e implementação em decisões automatizadas. A explicabilidade deve ser compreendida ainda como uma característica que se refere não apenas aos titulares de dados, por conta dos direitos afetados por tais decisões, mas também para profissionais que interagem com tais aplicações, ou que as utilizam como subsídios para as suas decisões.

Já apresentamos as diversas formas de explicação e diferentes nomenclaturas encontradas na literatura sobre sistemas computacionais nas seções 2.1.2 e 4.2.1.3, onde discutimos como as explicações podem ser locais, voltadas à compreensão do resultado de determinada aplicação, ou globais, voltadas à compreensão do processo pelo qual o sistema atinge os resultados[623]. Podem ainda se diferenciar entre explicações causais, que se relacionam aos elementos constitutivos de determinada conclusão, ou contrafactuais, referindo-se a exemplos hipotéticos que fornecem juízos alternativos pelos quais se pode compreender a lógica de uma decisão[624].

Outra diferenciação importante sobre os tipos de explicação diz respeito a sua natureza direta, baseada na descrição sobre o funcionamento do modelo, ou

[623] GUIDOTTI, R. *et al.* Local Rule-Based Explanations of Black Box Decision Systems. *ArXiv*, 2018. Disponível em: http://arxiv.org/abs/1805.10820. Acesso em: 30 jun. 2020.
[624] WACHTER, S.; MITTELSTADT, B.; RUSSELL, C. Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR. *SSRN Electronic Journal*, v. 31, n. 2, 2017. Disponível em: https://www.ssrn.com/abstract=3063289. Acesso em: 27 maio. 2020.

agnóstica, focando na relação entre os dados de entrada e de saída, sem referência ao modelo e aos processos intermediários. Neste ínterim pode-se realizar explicações que combinem esses diferentes elementos do ponto de vista de sua forma e que, além disso, apresentem diferentes conteúdos, como as fontes de dados utilizadas, os métodos empregados, as bases legais para o processamento, sobre a segurança, precisão e acurácia do modelo etc.

Tendo em vista os contextos de processamentos automatizados e os fundamentos apresentados neste trabalho para o direito à explicação como corolário da autodeterminação informacional e do devido processo informacional, a nossa proposta centrada no sujeito de direitos leva em consideração quatro dimensões de interesse: (i) o contexto de interação, (ii) as necessidades do sujeito, (iii) a complexidade do sistema utilizado na tomada de decisão e (iv) a possibilidade de desafiar a decisão.

Frente a diversidade de modelos de explicação é importante que as organizações se dediquem na escolha das estratégias para a implementação desse direito à explicação considerando o sujeito de direitos em várias dimensões. A partir de um diagnóstico inicial abrangente apresentado na seção anterior, é possível identificar uma série de instrumentos que podem auxiliar na identificação dos riscos e da melhor escolha para a implementação. Os instrumentos jurídicos e técnicos são relevantes em todas as dimensões e etapas do desenvolvimento de aplicações. No entanto, como veremos, para algumas dessas dimensões, o mais relevante são aspectos relacionados à proteção de dados e mitigação de impactos, em outros momentos teremos a proeminência de ferramentas de técnicas e de *design*.

A dimensão, o contexto da interação, envolve o reconhecimento sobre o domínio onde a aplicação está inserida, sobre os riscos, impactos e direitos que podem ser afetados pelo processamento. Nessa avaliação, deve-se considerar o papel da organização na cadeia de processamento, seja internamente em relação a sua governança, seja externamente em relação aos seus parceiros. Para cada tipo de tratamento, deve-se considerar as bases legais para o tratamento de dados, os ordenamentos e legislações aplicáveis ao domínio de atuação. É preciso considerar ainda a forma de interação, visto que o sujeito de direitos pode manter interações com pessoas ou com sistemas automatizados. Outro aspecto relevante diz respeito à urgência ou o tempo disponível na decisão, relacionado a sua reversibilidade ou grau de impacto.

O contexto da decisão automatizada, que deve orientar o desenvolvimento da explicação, pode ser compreendido a partir de uma série de instrumentos. Alguns deles previstos na legislação e já apresentados neste trabalho, como os Relatórios de Impacto e Proteção de Dados. A depender dos riscos envolvidos no domínio de aplicação ou da própria aplicação, deve-se considerar a realização de outras avaliações, inseridas nos relatórios de impactos ou produzidas em relatórios separados, como os discutidos na seção 5.1.4, por exemplo, relatórios de impacto aos direitos humanos ou relatórios de impacto algorítmico.

Uma aplicação que trate dados de forma massiva e promova inferências sobre a os sujeitos para realizar com base nisso tratamentos diferenciais no fornecimento de serviços ou bens, como as plataformas de prestação de serviços mencionadas no caso número B da seção anterior, por exemplo, apresentam um contexto definido. Pelo tratamento massivo de dados, há uma relação proeminente de cunho econômico com uma assimetria informacional em favor das plataformas e há além disso uma série de elementos que podem incorrer em relações jurídicas de cunho trabalhista de forma que um relatório de proteção de dados torna-se recomendável do ponto de vista da LGPD.

No entanto, já apontamos como o uso de algoritmos implica em questões que vão além da proteção de dados pessoais dos sujeitos individuais, trazendo efeitos sistêmicos. O uso de incentivos ou tarifas diferenciais em decorrência de horário, por um exemplo, podem incentivar motoristas a realizarem jornadas muito longas, de forma que a plataforma da Uber, por exemplo, agiu por limitar a jornada[625]. Estudos demonstram como o próprio sistema de classificação baseado na avaliação dos usuários pode resultar em efeitos discriminatórios[626]. Nesse sentido, outras avaliações para medir os impactos algorítmicos ou mesmo relatórios de impacto aos direitos humanos podem se mostrar necessárias.

A segunda dimensão a se considerar na implementação do direito à explicação diz respeito às necessidades dos usuários. Sob a denominação das necessidades do sujeito, podemos compreendê-las ao menos sob duas perspectivas. A primeira delas diz respeito à dimensão imediata da informação. Frente a uma decisão automatizada

---

[625] INTRODUCING A NEW FEATURE: DRIVING HOURS LIMIT. *UBER*, 2018. Disponível em: https://www.uber.com/he/blog/driving-hours-limit/. Acesso em: 5 jun. 2021.

[626] ROSENBLAT, A. *et al.* Discriminating Tastes: Uber's Customer Ratings as Vehicles for Workplace Discrimination. *Policy & Internet,* [*S. l.*], v. 9, n. 3, p. 256–279, 2017. Disponível em: https://doi.org/10.1002/poi3.153. Acesso em: 21 jun. 2021.

que foi ou será realizada, deparamo-nos com a questão de quais são as informações necessárias e que deveriam constar em uma explicação. Essa não é uma questão simples, visto que há tantas explicações quanto sejam possíveis para fenômenos, desde as mais simples às mais complexas. Neste sentido, a explicação mais correta e mais completa possível nem sempre pode se demonstrar a mais útil e corresponder a expectativa desse sujeito.

Muitos autores têm apontado que não há um conceito único sobre a explicação. No entanto, a partir da compreensão da autodeterminação informativa e do devido processo informacional, a nossa proposta se alinha àquelas perspectivas de que esta deve ser centrada no sujeito, sob um ponto de vista pragmático, em relação à possibilidade de exercício dos seus direitos[627]. A depender da aplicação, o tipo de informação necessária para que se compreenda uma decisão pode ser diverso. Em alguns casos pode ser necessário explicar o próprio direito envolvido, em alguns deles, a natureza da decisão automatizada, em outros, sua segurança ou precisão, a fonte dos dados utilizados no modelo.

Nesse sentido, é preciso depreender uma série de informações relevantes para a interpretação da decisão. Essas informações são depreendidas a partir dos diferentes instrumentos de avaliação disponíveis como os relatórios de impacto, os princípios da proteção de dados e das demais legislações aplicáveis ao caso.

A título de exemplo, imaginemos uma aplicação de análise antifraude apresentada na seção anterior. Apontamos como duas das principais causas de erros neste tipo de aplicação consistem em erros derivados dos critérios utilizados na programação ou das bases de dados utilizadas. Apontamos como tal decisão pode causar um impacto significativo. Como etapa de um processo do direito administrativo, argumentamos como tais decisões precisam constar dos princípios do processo administrativo, na qual devem estar presentes os fundamentos de fato e de direito que determinaram a decisão, conforme o art. 2º, inciso II, da Lei nº 9.784 de 1999, a “Lei do Processo Administrativo”.

Isso posto, a pessoa sujeita a um bloqueio ou notificação resultando de tal sistema necessita conhecer qual é o direito envolvido no benefício, a legislação que o

[627] WATSON, D. S.; FLORIDI, L. The explanation game: a formal framework for interpretable machine learning. *Synthese*, 2020. Disponível em: https://doi.org/10.1007/s11229-020-02629-9. Acesso em: 8 jun. 2020; e DOSHI-VELEZ, F.; KIM, B. Towards A Rigorous Science of Interpretable Machine Learning. *ArXiv [cs, stat*], [*S. l.*], 2017. Disponível em: http://arxiv.org/abs/1702.08608. Acesso em: 30 jun. 2020.

regula, os critérios da concessão de benefícios, qual o resultado da decisão, porque seu perfil foi considerado inadequado e, neste caso, diante da possibilidade clara de danos a seus direitos, quais são as possibilidades de que tal sistema tenha apresentado o erro, bem como as formas de corrigi-lo. Por certo, nas aplicações, as avaliações apontaram erros mais específicos e outras questões relevantes que também devem ser consideradas.

Mas, além desse sentido informacional, do encadeamento lógico de informações necessárias para a compreensão da decisão, demonstramos como a explicação deve ser compreendida como um processo dialógico. Nesse sentido, outro aspecto relevante ao se considerar as necessidades dos sujeitos de direitos das decisões, além da mera explanação do fenômeno, é a sua efetiva compreensão[628].

Nesse sentido, a implementação do direito à explicação depende do reconhecimento dos sujeitos a quem se direcionam as possíveis explicações.[629] Devem ser conhecidas as pessoas concretas que estarão interagindo com a aplicação e submetidos à decisão com suas características cognitivas, incluindo possíveis limitações. Devem ser compreendidas em seus aspectos sociais e culturais. No nosso exemplo de um sistema antifraude a concessão de benefícios sociais, a condição socioeconômica, por exemplo, deve ser levada em consideração. A depender da idade da população a qual essa política se destina, outras questões certamente irão emergir.

Discutimos na seção 4.2.1 como a capacidade de compreensão sobre processos computacionais está desigualmente distribuída entre a população. E, além disso, discutimos como os indivíduos, nas suas interações cotidianas, realizam explicações, de forma que a compreensão cotidiana ocorre de forma bem diversa dos processos técnicos de conhecimento. Um trabalho seminal de Tim Miller, baseando-se em uma revisão do tema nas ciências sociais, demonstra como a) as explicações que os indivíduos costumam realizar cotidianamente sobre um fenômeno são contrastivas em relação a outros fatos que poderiam resultar em eventos diferentes; b) que as explicações são seletivas em função das informações mais relevantes e

---

628 BOHLENDER, D.; KÖHL, M. A. Towards a Characterization of Explainable Systems. *ArXiv* [cs], [*S. l.*], 2019. Disponível em: http://arxiv.org/abs/1902.03096. Acesso em: 6 out. 2020.

629 É importante destacar que as explicações podem ser endereçadas não apenas aos titulares dos dados, mas aos agentes públicos, organizações da sociedade civil etc, principalmente em aplicações cujos efeitos indesejados possam incluir efeitos sistêmicos. Contudo, a interlocução com essas entidades pauta-se por uma abordagem institucional, diferente da abordagem centrada no sujeito de direitos afetado pelas decisões.

mais importantes envolvidas na pergunta do agente receptor; c) que as probabilidades não são suficientes para explicações, a menos que sejam acompanhadas de mais explicações causais; e d) que as explicações são processos sociais e derivam de procedimentos dialógicos, relacionais, num compartilhamento de conhecimento[630].

O trabalho de Miller aponta caminhos relevantes com base na revisão dos estudos da sociologia e da psicologia social sobre como as pessoas avaliam as explicações no cotidiano. A qualidade da explicação é avaliada com base na probabilidade de estar correta, na sua simplicidade, na generalidade para tratar do problema e na coerência com as crenças anteriores[631]. Nesse sentido, explicações não devem ser compreendidas como encadeamentos de associações entre fatos, devem ser compreendidas como relações contextuais e dialógicas. O autor argumenta que, enquanto pode haver várias explicações possíveis para um fenômeno, é no contexto da interação social que se deve descobrir a melhor delas.

A instrumentalização dessa reflexão sobre a efetividade da explicação depende de um esforço de *design*. Como o contexto das interações para muitas atividades de decisões automatizadas consiste em relações entre o sujeito e uma interface eletrônica, o conceito de *human computer interactions*[632], comum nos estudos sobre desenvolvimento de *softwares*, pode servir como um modelo para a implementação do direito à explicação. Apresentamos essa discussão na seção 5.1.5 ao tratarmos do conceito de *explainability by design*.

Demonstramos como uma categoria central para a concepção de sistemas explicáveis é justamente a definição de um grupo alvo para a explicação. Embora não seja possível definir precisamente um grupo e que a mesma aplicação pode incidir sobre uma diversidade de sujeitos materialmente muito diferentes, as explicações que não tenham intenção de realizar generalizações e abstrações são pouco úteis, a menos que pensemos em sistemas completamente artesanais, onde uma explicação seja construída para cada pessoa.

Nesse sentido, esse grupo alvo pode ser pensado a partir de agentes representativos que possuem determinadas características, consideradas no

---

630 MILLER, T. Explanation in Artificial Intelligence: Insights from the Social Sciences. *ArXiv [cs]*, [*S. l.*], 2018. Disponível em: http://arxiv.org/abs/1706.07269. Acesso em: 14 ago. 2020.

631 MILLER, T. Explanation in Artificial Intelligence: Insights from the Social Sciences. *ArXiv [cs]*, [*S. l.*], 2018. Disponível em: http://arxiv.org/abs/1706.07269. Acesso em: 14 ago. 2020.

632 SINHA, G.; SHAHI, R.; SHANKAR, M. Human Computer Interaction. *In*: INTERNATIONAL CONFERENCE ON EMERGING TRENDS IN ENGINEERING AND TECHNOLOGY, 3., 2010. p. 1–4. Disponível em: https://doi.org/10.1109/ICETET.2010.85. Acesso em: 21 jun. 2021.

desenvolvimento da aplicação como os possíveis sujeitos cujos direitos serão afetados pelas aplicações. É possível que tal concepção de grupo seja derivada de técnicas e métodos de estudos antropológicos e sociológicos. É possível, ainda, que setores da organização possuam conhecimento sobre tais sujeitos, de forma que essa concepção pode ser construída a partir de diferentes pontos de vista.

Quando se discutiu o conceito de *explainability by design*, a proposta de avaliação da explicabilidade de sistemas e de como os estudos experimentais com usuários sugerem que essas explicações poderiam se basear em categorias comuns da literatura de *Human Computer Interaction* (HCI): a confiança, a *human-likeness*, a justificação adequada e a compreensibilidade[633]. Outros trabalhos apontam como se utilizar metodologias empíricas para identificar as melhores explicações e destacam a importância do desenvolvimento da interdisciplinaridade e da sensibilização da comunidade de desenvolvedores e cientistas de dados[634] .

Outra dimensão central a ser considerada na implementação do direito à explicação diz respeito à complexidade do sistema utilizado na tomada de decisão. Discutimos na seção 4.2.1 como a complexidade dos sistemas algorítmicos impõe determinadas condições para sua apreensão pelos sujeitos. Nesse sentido, quanto maior a complexidade, a depender da metodologia empregada, a interpretabilidade do sistema pode ser maior ou menor.

A depender da natureza dos direitos afetados na decisão automatizada, é preciso considerar se o grau de opacidade, de taxas de erros e acurácia do modelo demandam cuidados adicionais de mitigação de risco e um acompanhamento mais atencioso, incluindo maior supervisão de seres humanos. Se algumas aplicações são pouco interpretáveis, até mesmo por especialistas e desenvolvedores, o exercício do direito à explicação encontra uma barreira adicional. Alguns trabalhos têm fornecido diversas abordagens para modelos de *machine learning* interpretáveis que podem ser úteis para considerarmos a implementação do direito à explicação[635].

Discutimos na seção 4.1.1 como a revelação de informações sobre o

---

633 EHSAN, U.; RIEDL, M. O. *On Design and Evaluation of Human-centered Explainable AI systems*. Glasgow, 2019. Disponível em: https://www.cc.gatech.edu/~riedl/pubs/ehsan-chi-hcml19.pdf. Acesso em: 22 set. 2020.

634 LIAO, Q. V.; GRUEN, D.; MILLER, S. Questioning the AI: Informing Design Practices for Explainable AI User Experiences. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems,* [*S. l.*], p. 1–15, abr. 2020. Disponível em: https://doi.org/10.1145/3313831.3376590. Acesso em; 21 jun. 2021.

635 MOLNAR, C. *Interpretable Machine Learning*. [*S. l.*]: [*s. n.*], 2020. *E-book*.

funcionamento de algoritmos podem representar um risco ao segredo comercial de algumas organizações e, nesse sentido, é comum que as organizações desenvolvam modelos complexos como forma de evitar o risco de perder o seu segredo de negócio. No entanto, tendo em vista as necessidades do sujeito e os direitos de transparência, que podem ser reivindicados perante as autoridades ou em juízo, é prudente que as organizações levem em conta a necessidade de que seus modelos sejam explicáveis de maneira satisfatória, como forma de evitar litigâncias em torno da revelação de seus códigos. Há diversas modalidades de explicação que são agnósticas em relação ao código, que podem servir para proteger o segredo de determinados processos, fornecendo, contudo, as informações essenciais para o exercício de direitos.

Uma última dimensão a ser considerada para a implementação do direito à explicação centrado no sujeito de direitos diz respeito à possibilidade de desafiar a decisão automatizada. Já discutimos como a possibilidade de revisão de decisões automatizadas previsto no art. 20 é um corolário da autodeterminação informacional e do devido processo informacional. É justamente vinculada a essa possibilidade que o direito à explicação surge como condição necessária[636] .

A implementação do direito à explicação, portanto, deve observar a disponibilidade de meios de revisão da decisão automatizada, seja por determinação legal, seja em relação às possibilidades técnicas disponíveis para sua aplicação, dos custos de implementação desse processo e do impacto oriundo de uma decisão malsucedida.

O direito de revisão de decisões automatizadas ainda é alvo de debates, como discutiu-se ao longo deste trabalho. Há críticas contundentes à redação do art. 20, por restringir demais o conceito de decisões automatizadas. Além disso, há críticas quanto à supressão da previsão de revisão humana do texto do art. 20 da LGPD de forma que a discussão em torno do tema seria suficiente para outra tese.

Após nossa exposição, contudo, e ao final desse trabalho, esperamos que se compreenda que o direito à explicação, enquanto corolário do devido processo informacional, extrapola a previsão do art. 20, de forma que mesmo que uma decisão seja tomada de forma apenas parcialmente automatizada, o não fornecimento de

---

[636] MULLIGAN, D. K.; KLUTTZ, D.; KOHLI, N. *Shaping Our Tools*: Contestability as a Means to Promote Responsible Algorithmic Decision Making in the Professions. Rochester, NY: Social Science Research Network, 2019. SSRN Scholarly Paper. Disponível em: https://doi.org/10.2139/ssrn.3311894. Acesso em: 14 jan. 2021.

informações relevantes e significativas, capazes de serem compreendidas pelo titular dos dados, representa uma violação grave de seus direitos. De forma que essa instrumentalização do direito à explicação deve ser compreendida de forma mais ampla do que uma mera obrigação de *compliance*. Nesse sentido, podemos acrescentar que apesar da retirada da previsão de revisão humana do texto da LGPD, é prudente observar como em alguns domínios e alguns contextos tal participação seja extremamente necessária.

O quadro abaixo apresenta uma visão esquemática sobre as dimensões da abordagem centrada no sujeito de direitos, os fatores relevantes para o desenvolvimento da estratégia e os instrumentos técnicos e jurídicos que podem ser utilizados.

QUADRO 2 – Dimensão, Fatores e Instrumentos técnicos e jurídicos

DIMENSÃO	FATORES	INSTRUMENTOS TÉCNICOS E JURÍDICOS
CONTEXTO	(i) Domínio de aplicação, riscos, impactos e direitos envolvidos na decisão. Papel da organização, cadeia de processamento de dados, bases jurídicas para o tratamento de dados, legislações aplicáveis, forma de interação. A urgência da decisão	Estudos diagnósticos LIA Relatórios de Impacto Auditorias
NECESSIDADES	(ii) expectativas dos sujeitos de direitos, relações jurídicas da decisão. Necessidade e finalidade da decisão, direitos envolvidos, dimensões éticas da decisão, informações relevantes para o exercício dos direitos.	
	(iii) público-alvo, forma de explicação, especificidade da matéria, conhecimentos necessários para compreensão	Relatórios de Impacto Estudos de casos XbD Participação Humana
COMPLEXIDADE DO SISTEMA	iv) discussão sobre a interpretabilidade do sistema, sobre o domínio e a finalidade com que foi desenvolvido, sobre o estado da arte do campo de aplicação técnico, sob os riscos ao modelo de negócios nos casos de revelações de informações sobre o sistema.	
POSSIBILIDADE DE DESAFIAR A DECISÃO	(v) procedimentos necessários para desafiar a decisão considerando o devido processo informacional e a autodeterminação informativa	Relatórios de Impacto LIA Auditorias Interface para questionamento da decisão Revisão Humana

Fonte: Elaboração própria

Na seção 5.1.5, apresentamos uma série de modelos de explicação, a partir da sistematização feita pelo ICO e pelo The Alan Turing Institute sobre as formas de explicação da inteligência artificial, que podem ser combinadas em estratégias efetivas a partir da consideração dos elementos aqui apresentados. Cabe apontar, a título de conclusão, que a explicabilidade e o direito à explicação não devem ser

compreendidos como um simples ato de fornecimento de informações. Antes de tudo, consistem em uma preocupação transversal a toda cadeia de tratamento de dados da qual a decisão automatizada faz parte.

5.2.2 Proposta de um *framework* de explicabilidade a partir da cláusula geral do devido processo informacional

Na seção 2.2.3 (“Direito à explicação como corolário do devido processo informacional”), analisamos o debate acadêmico norte-americano em torno da cláusula geral do devido processo informacional e como essa noção vem sendo gradualmente incorporada ao debate acadêmico nacional e à jurisprudência constitucional brasileira. Na presente seção, propomos um *framework* de explicabilidade elaborado a partir da cláusula geral do devido processo informacional, que, sem a pretensão de ser exaustivo, absoluto ou aplicável diretamente à generalidade de casos, possui função orientativa, estando estruturado em torno de princípios, de diretrizes e de eixos balizadores da construção de uma explicação minimamente satisfatória e adequada.

No subtópico 5.2.1, pontuamos que um modelo de explicabilidade deve partir de ao menos três diretrizes ou considerações preliminares: primeiro, que um *framework* de explicabilidade, longe de ser um modelo fixo, genérico ou absoluto, deve ser compreendido como um conjunto de ferramentas técnico-jurídicas que, ao serem articuladas, tornam-se úteis à consecução de diferentes objetivos; segundo, que um modelo de explicabilidade deve ser concebido de forma contextual, aplicando os instrumentos da “caixa de ferramentas” do *framework* à luz dos diferentes contextos e riscos de cada atividade considerada; e terceiro, que a instrumentalização do direito à explicação deve se dar a partir de um olhar centrado no sujeito de direitos e suas necessidades.

Partindo dessas três considerações preliminares, propomos um *framework* de explicabilidade concebido em torno de um elemento estruturante e cinco qualificadoras da cláusula geral do devido processo informacional. Assim como a cláusula geral do devido processo legal encontra-se pautada na ideia de prevenção e de observância aos ditames do Estado Democrático de Direito, visando, assim, garantir que o Estado se valha de meios adequados, razoáveis e proporcionais para evitar abusos de direito e violação a direitos fundamentais, também a cláusula geral

do devido processo informacional deve ser pautada pela ideia geral de prevenção, sobretudo por dizer respeito a um objeto cada vez mais regulado a partir da noção de risco e prevenção de danos. O princípio da prevenção, portanto, é tomado como elemento estruturante e transversal a todo o *framework* de explicabilidade ora proposto.

A este elemento estruturante somam-se cinco qualificadoras, a saber: (i) isenção; (ii) informação; (iii) compreensão; (iv) recorribilidade; e (v) revisão. Embora não haja clareza sobre quais seriam os elementos conformadores da cláusula geral do devido processo legal[637] e do devido processo informacional[638], acreditamos que estas cinco qualificadoras nos permitem reconhecer os contornos e a espinha dorsal destas cláusulas "guarda-chuva", permitindo estabelecer uma aproximação entre ambos os contextos considerados (decisões privadas/estatais e decisões automatizadas).

Ao longo desta seção, analisaremos cada uma dessas qualificadoras sob a ótica tanto do devido processo legal quanto do devido processo informacional, contextualizando-as nos processos automatizados que levam à elaboração de decisões com impactos para os sujeitos de direitos. Após a apresentação destas cinco qualificadoras, cuidaremos mais especificamente de duas delas (informação e compreensão), por entendermos estarem mais diretamente relacionadas à construção

---

[637] Conforme assinala Arruda, a doutrina estrangeira tem ressaltado a amplitude desta cláusula geral, que se revela como um conceito bastante amplo e indefinido, abarcando uma série de liberdades e garantias fundamentais do indivíduo em suas relações verticais e horizontais. De acordo com a autora, e conforme destacado nos Capítulos 2 e 3, no contexto norte-americano, a cláusula geral do devido processo legal tem sua origem em 1789 com a promulgação da 5ª e 14ª emendas à Constituição, "[...] na carta denominada Bill of Rights, onde se declaravam os direitos dos cidadãos em face do poder do governo, estabelecendo garantias expressas como o due process of law, princípio segundo o qual nenhuma pessoa seria privada de sua vida, liberdade ou propriedade, ou seja, dos seus direitos fundamentais, sem o devido processo legal." (ARRUDA, C. S. L. de. Breve estudo hermenêutico-epistemológico da cláusula do "devido processo legal". *Revista CEJ*, Brasília, Ano XXI, n. 73, set./dez. 2017. p. 55). No Brasil, a Constituição Federal de 1988, claramente informada pelo texto da Constituição norte-americana, igualmente previu uma cláusula bastante ampla garantindo o devido processo legal quando estiverem em risco o patrimônio ou as liberdades do cidadão (CF 1988, art. 5º, caput e LIV). Pela própria amplitude da cláusula, nela encontram-se incluídas todas as possíveis garantias destinadas a salvaguardar as liberdades e a propriedade dos cidadãos, sendo impraticável mensurar, em rol taxativo, todos os elementos que a integram, não tendo o presente trabalho tal pretensão.

[638] Assim como o devido processo legal, a cláusula geral do devido processo informacional possui escopo amplo e contornos pouco claros, abarcando uma série de garantias voltadas à proteção dos direitos e liberdades dos indivíduos no contexto de decisões automatizadas. Ao enunciarmos as cinco qualificadoras que conformam a cláusula geral do devido processo informacional para os fins deste trabalho não pretendemos conferir-lhe uma definição exaustiva e acabada. As cinco qualificadoras foram extraídas da literatura sobre devido processo informacional e buscam em alguma medida estabelecer certo paralelo com algumas das garantias do devido processo legal tal como tradicionalmente concebido pela dogmática jurídica.

de uma explicação. Nessa fase, buscaremos analisar como essas duas qualificadoras podem ser garantidas, contextualizando-as em relação às características e necessidades dos sujeitos de direitos, a partir de uma perspectiva *ex ante* e *ex post*, considerando a caixa de ferramentas técnico-jurídicas disponível. Neste ponto, cumpre esclarecer que, embora as três qualificadoras restantes (isenção, recorribilidade e revisão) sejam igualmente relevantes, não constitui objetivo deste trabalho oferecer um *framework* de revisão de decisões automatizadas ou mesmo orientações sobre como garantir que sistemas automatizados sejam livres de vieses, aspectos que fogem ao escopo desta pesquisa.

A primeira qualificadora, portanto, consiste na garantia de isenção do processo. No devido processo legal, esta garantia se traduz no direito do cidadão ter sua violação de direitos analisada por um ente neutro e isento do Estado, como, por exemplo, um juiz. De acordo com Cintra, Dinamarco e Grinover, a imparcialidade do órgão de jurisdição é condição intrínseca para que este possa exercer sua função dentro do processo, sendo ainda pressuposto de validade da relação processual. Nesse sentido, o juiz deve colocar-se simetricamente entre as partes e acima delas, em posição de imparcialidade. Uma vez constatada a incapacidade subjetiva do juiz, ou seja, uma vez comprometida sua posição de isenção, a relação processual resta profundamente afetada. É por esta razão que o direito constitucional procurou cercar o juiz e o jurisdicionado de certos direitos e garantias, como é o caso das suspeições e impedimentos, o princípio do juiz natural, as vedações aos tribunais de exceção, dentre outros mecanismos de salvaguarda processual.[639] Em suma, de acordo com os autores:

> A imparcialidade do juiz é uma garantia de justiça para as partes. Por isso, têm elas o direito de exigir um juiz imparcial: e o Estado, que reservou para si o exercício da função jurisdicional, tem o correspondente dever de agir com imparcialidade na solução das causas que lhe são submetidas.[640]

Transposta para a lógica do devido processo informacional, é possível traduzir esta garantia como o direito do titular de dados pessoais se sujeitar a um processo automatizado, neutro e isento, como, por exemplo, sem vieses discriminatórios. O

[639] CINTRA, A. C. de A.; GRINOVER, A. P.; DINAMARCO, C. R. *Teoria Geral do Processo*. 27. ed. São Paulo: Malheiros, 2011. p. 58.
[640] CINTRA, A. C. de A.; GRINOVER, A. P.; DINAMARCO, C. R. *Teoria Geral do Processo*. 27. ed. São Paulo: Malheiros, 2011. p. 59.

emblemático caso do COMPAS, sistema automatizado largamente empregado pela justiça norte-americana para embasar análises de risco de reincidência e o cálculo de pena de réus da justiça criminal, é particularmente interessante para ilustrar a importância e os riscos do comprometimento da isenção em um processo decisório automatizado. Larson, Mattu, Kirchner e Angwin conduziram um extenso estudo com o objetivo de analisar a acurácia do algoritmo de cálculo de reincidência e verificar se seu funcionamento estaria enviesado em relação a determinados grupos. As conclusões do estudo apontaram para a existência de um forte viés racial e discriminatório no algoritmo de análise de reincidência do COMPAS, que chegou a classificar indivíduos negros como possuindo duas vezes maior risco de reincidência em relação a indivíduos brancos, mesmo que não viessem efetivamente a reincidir, e, inversamente, tendia a classificar indivíduos brancos como sendo de baixo risco, mesmo que viessem efetivamente a reincidir.[641] No Brasil, o emprego de tecnologias de reconhecimento facial para fins de persecução penal e segurança pública têm avançado vertiginosamente nos últimos anos, levantando preocupações entre pesquisadores e sociedade civil organizada quanto ao seu potencial discriminatório sobre a população racializada, que historicamente tende a ser estigmatizada e discriminada pelo aparato de segurança e justiça estatal.[642] Um estudo realizado pela Rede de Observatórios de Segurança é exemplificativo deste cenário: a instituição monitorou a implementação de tecnologias de reconhecimento facial para fins de segurança pública e persecução penal em cinco estados brasileiros ao longo do ano de 2019 e identificou que, dos 151 casos monitorados, 90,5% dos presos por reconhecimento facial no Brasil eram negros.[643]

Assim como no devido processo legal, portanto, a isenção revela-se como uma condição para a garantia da justiça procedimental em uma relação processual mediada por algoritmos. Um devido processo informacional requer, assim, que os

---

641 LARSON, J.; MATTU, S.; KIRCHNER, L.; ANGWIN, J. How We Analyzed the COMPAS Recidivism Algorithm. *ProPublica*, 23 maio 2016. Disponível em: https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm. Acesso em: 31 maio 2021; e SPIELKAMP, M. Inspecting Algorithms for Bias. *MIT Technology Review*, 12 jun. 2017. Disponível em: https://www.technologyreview.com/2017/06/12/105804/inspecting-algorithms-for-bias/. Acesso em: 31 maio 2021.

642 SILVA, T. Reconhecimento facial deve ser banido: veja dez razões. *Tarcízio Silva*, 16 maio 2021. Disponível em: https://tarciziosilva.com.br/blog/reconhecimento-facial-deve-ser-banido-aqui-estao-dez-razoes/. Acesso em: 31 maio 2021.

643 NUNES, Pablo. Levantamento revela que 90,5% dos presos por monitoramento facial no Brasil são negros. *The Intercept*, 21 nov. 2019. Disponível em: https://theintercept.com/2019/11/21/presos-monitoramento-facial-brasil-negros/. Acesso em: 31 maio 2021.

algoritmos mediadores da relação, enquanto instâncias responsáveis pela atribuição de bens e direitos, operem de forma justa, isenta, imparcial e livre de vieses discriminatórios, garantindo igualdade de tratamento aos destinatários da decisão.

A segunda qualificadora consiste na garantia de informação acerca do processo. No contexto do devido processo legal, essa garantia se traduz como o direito do jurisdicionado ter acesso às informações sobre a relação processual, tais como os autos de um processo judicial, que são acessíveis às partes e, majoritariamente, ao público para cumprir com as obrigações de transparência. Essa garantia se expressa, por exemplo, através do princípio da publicidade do processo, que visa cercar a própria atividade do Estado juiz de algum grau de controle social e legitimidade ao submetê-la ao escrutínio público (publicidade popular), ressalvadas as hipóteses em que prevaleçam o sigilo e a intimidade das partes, ocasião em que os atos processuais são tornados públicos apenas com relação às partes e seus defensores (publicidade restrita).[644]

Na Constituição de 1988, essa regra geral de publicidade do processo encontra-se prevista no art. 5º, LX, que estabelece que a lei só poderá restringir a publicidade dos atos processuais quando a defesa da intimidade ou o interesse social o exigirem, e no art. 93, IX, que dispõe que todos os julgamentos dos órgãos do Poder Judiciário serão públicos, e fundamentadas todas as decisões, sob pena de nulidade, podendo a lei limitar a presença, em determinados atos, às próprias partes e a seus advogados, ou somente a estes, em casos nos quais a preservação do direito à intimidade do interessado no sigilo não prejudique o interesse público à informação. Neste aspecto, a garantia da informação confere uma dimensão pública ao processo, cercando-o de transparência e submetendo-o ao escrutínio social. Em relação às partes, mais especificamente, a garantia de informação pode ser extraída do princípio do contraditório e da ampla defesa, previsto no art. 5º, LV, da Constituição Federal, que assegura aos litigantes, em processo judicial ou administrativo, e aos acusados em geral o contraditório e ampla defesa, com os meios e recursos a ela inerentes. Conforme assinalam Cintra *et al.*[645], decorre do contraditório e da ampla defesa a necessidade de garantir aos litigantes ciência sobre os atos praticados pelo juiz e pela

[644] CINTRA, A. C. de A.; GRINOVER, A. P.; DINAMARCO, C. R. *Teoria Geral do Processo*. 27. ed. São Paulo: Malheiros, 2011. p. 75.
[645] CINTRA, A. C. de A.; GRINOVER, A. P.; DINAMARCO, C. R. *Teoria Geral do Processo*. 27. ed. São Paulo: Malheiros, 2011. p. 62.

contraparte, condição necessária para o próprio exercício do contraditório. De acordo com os autores, entre nós, dentre outros instrumentos, a ciência dos atos processuais se dá mediante a citação, a intimação e a notificação.[646]

Um exemplo de como essa qualificadora pode ser operacionalizada no contexto do devido processo informacional consiste no julgamento sobre o emprego de metodologias de pontuação de crédito realizadas pelo Superior Tribunal de Justiça em 2014. Nos âmbitos do Recurso Especial nº 1.419.697/RS[647] e do Recurso Especial nº 1.457.199/RS[648], de relatoria do Min. Paulo de Tarso Sanseverino, ambos analisados no tópico 3.2.1 deste trabalho, reconheceu-se a legalidade de metodologias de análise de risco de crédito à luz das disposições do Código de Defesa do Consumidor e da Lei do Cadastro Positivo, desde que observadas determinadas garantias legais. No conjunto de cinco teses fixadas no julgamento dos referidos recursos repetitivos, aplicáveis nos casos envolvendo o tratamento de dados para fins de análise de risco de crédito, destacam-se as amplas garantias de informação e transparência conferidas. Nos termos da tese n. 3, na avaliação do risco de crédito "[...] devem ser respeitados os limites estabelecidos pelo sistema de proteção do consumidor no sentido da tutela da privacidade e da *máxima transparência nas relações negociais*, conforme previsão do CDC e da Lei n. 12.414/2011." (grifo nosso). A tese n. 4, por sua vez, estabelece que, não obstante a prescindibilidade de obtenção de consentimento do consumidor consultado, "devem ser a ele fornecidos *esclarecimentos, caso solicitados, acerca das fontes dos dados considerados (histórico de crédito), bem como as informações pessoais valoradas*." (grifo nosso).

Tanto no devido processo legal quanto no devido processo informacional, portanto, cercar o processo de mecanismos de transparência e garantir o direito de acesso do indivíduo a informações sobre a relação processual servem ao objetivo de garantir certa regularidade procedimental às decisões. Todavia, a garantia do mero acesso aos dados e às informações que serviram de substrato para a construção da decisão não necessariamente permite compreender como o processo funciona e

---

[646] CINTRA, A. C. de A.; GRINOVER, A. P.; DINAMARCO, C. R. *Teoria Geral do Processo*. 27. ed. São Paulo: Malheiros, 2011. p. 62.

[647] BRASIL. Superior Tribunal de Justiça. *Recurso Especial nº 1.419.697/RS*. Relator: Min. Paulo de Tarso Sanseverino. Julgado em: 12/11/2014. Data de Publicação: 17/11/2014. Disponível em: https://bdjur.stj.jus.br/jspui/bitstream/2011/114173/REsp1419697.pdf. Acesso em: 8 abr. 2021.

[648] BRASIL. Superior Tribunal de Justiça. *Recurso Especial nº 1.457.199/RS*. Relator: Min. Paulo de Tarso Sanseverino. Julgado em: 12/11/2014. Data de Publicação: 17/12/2014. Disponível em: https://bdjur.stj.jus.br/jspui/bitstream/2011/114173/REsp1457199.pdf. Acesso em: 8 abr. 2021.

como estes elementos são valorados, articulados e mobilizados para a tomada de uma decisão automatizada, pelo que se faz necessária a terceira qualificadora da cláusula geral do devido processo informacional.

Avançando em relação à qualificadora da informação, faz-se necessário ainda garantir que o titular efetivamente compreenda como determinada decisão automatizada foi alcançada, para além de garantir-lhe o mero acesso aos elementos (dados e informações) que embasaram a tomada da decisão. Analisada sob a ótica do devido processo legal, a qualificadora da compreensão consiste no direito do jurisdicionado de entender as razões, os fundamentos e o racional de uma decisão que lhe afeta, o que se garante, por exemplo, através dos deveres atribuídos ao Estado-juiz de expressamente motivar a sua decisão judicial, informando quais provas e argumentos foram aceitos, a valoração destes, os motivos pelos quais os demais argumentos e provas não foram acolhidos e o racional que levou à tomada da decisão.

No contexto do devido processo legal, portanto, a qualificadora da compreensão está intimamente relacionada à exigência de motivação das decisões judiciais. Conforme apontam Cintra *et al.*[649], tradicionalmente, o dever de motivação das decisões costumava ser entendido tão somente como uma das garantias processuais das partes, sendo necessário para que essas compreendessem efetivamente como a decisão fora alcançada e pudessem exercer o direito de impugnação com o objetivo de reformar a decisão judicial. Contemporaneamente, contudo, entende-se que esta garantia transcende a própria relação entre as partes e o Estado-juiz, que deixam de ser os seus únicos destinatários, assumindo uma função política e um destinatário mais amplo e difuso, prestando-se a garantir um maior nível de escrutínio público sobre as decisões judiciais, de modo a "[...] aferir-se em concreto a imparcialidade do juiz e a legalidade e justiça das decisões." Foi justamente em razão de sua função política que diversas constituições ergueram esta garantia ao *status* de norma constitucional. A Constituição de 1988, ao contrário do texto que a precedera, adotou expressamente esta garantia, ao dispor em seu art. 93, IX, que todas as decisões emanadas de órgãos do Judiciário deverão ser fundamentadas.

Aplicada ao contexto de decisões automatizadas, a qualificadora da compreensão pode ser entendida como o direito do titular obter uma explicação clara,

---

649 BRASIL. Superior Tribunal de Justiça. *Recurso Especial nº 1.457.199/RS*. Relator: Min. Paulo de Tarso Sanseverino. Julgado em: 12/11/2014. Data de Publicação: 17/12/2014. Disponível em: https://bdjur.stj.jus.br/jspui/bitstream/2011/114173/REsp1457199.pdf. Acesso em: 8 abr. 2021. p. 74.

suficiente e adequada de como funciona uma decisão automatizada e de como esta pode impactar seus direitos e liberdades fundamentais, dando-lhe condições para que possa contestá-la. Neste sentido, não seria suficiente o mero acesso aos dados utilizados e o fornecimento de informações genéricas, havendo a necessidade de explicitação de como os dados foram efetivamente tratados e valorados e do racional da lógica de funcionamento dos sistemas automatizados aos quais foram submetidos. Conforme abordamos no Capítulo 2, há diversas formas de garantir que o titular compreenda um processo automatizado de decisão ao qual esteja ou venha a estar sujeito, como, por exemplo, através de mecanismos *ex ante* de proteção, como documentação de trilhas de auditoria algorítmica, elaboração de relatórios de impacto a proteção de dados pessoais e realização de auditorias algorítmicas pelas autoridades de supervisão, bem como através de garantias *ex post*, como direitos de acesso mais robustos envolvendo explicações globais e locais, ou seja, informações claras, adequadas e significativas tanto sobre o funcionamento dos sistemas, incluindo análise de códigos fonte, como também sobre os aspectos e fundamentos de uma decisão individualmente considerada.

Um exemplo prático de como essa qualificadora do devido processo informacional pode ser operacionalizada encontra-se na Ação Trabalhista nº 0000335-45.2020.5.09.0130, movida em face da 99 Táxi, na qual, para verificar a existência de relação de emprego, o juiz determinou a realização de auditoria algorítmica de modo a esclarecer, dentre outros aspectos, (i) as condições em que as chamadas eram distribuídas pelo aplicativo, (ii) a forma em que se estabelecem os valores a serem repassados e (iii) a existência de condições ou preferências no acesso e na distribuição dos chamados em função da avaliação do motorista ou do número de aceites de outras corridas, conforme analisamos no item 4.3.3 deste trabalho. Mais do que requisitar o mero acesso ao código do aplicativo ou aos dados utilizados, requereu-se uma explicação sobre o funcionamento global da aplicação destinada a subsidiar o livre convencimento do juiz e a defesa da parte autora.

A quarta qualificadora da cláusula geral do devido processo informacional consiste na garantia de recorribilidade, ou seja, na possibilidade de contestar a decisão automatizada. No contexto do devido processo legal, essa garantia se expressa através do princípio do duplo grau de jurisdição. De acordo com Cintra *et*

al.[650], este princípio garante a possibilidade de revisão de uma causa já julgada pelo juízo de primeiro grau, pela via recursal, por uma instância de segundo grau. Conforme apontam os autores, o "[...] princípio do duplo grau de jurisdição funda-se na possibilidade de a decisão de primeiro grau ser injusta ou errada, daí decorrendo a necessidade de permitir sua reforma em grau de recurso", dando ao ofendido uma oportunidade de reexame da sentença que o afetou.[651] Compreendida a partir da lógica do devido processo informacional, essa garantia se traduz no direito do titular de solicitar a revisão de decisões tomadas unicamente com base em tratamento automatizado de dados pessoais que afetem seus interesses, incluídas as decisões destinadas a definir o seu perfil pessoal, profissional, de consumo e de crédito ou os aspectos de sua personalidade, nos termos do art. 20 da LGPD. Nesse sentido, a quarta qualificadora possui como requisito as duas anteriores, ou seja, devem ser garantidas a informação e a compreensão, ou, em outras palavras, um efetivo direito à explicação, para que o titular seja capaz de contestar ou recorrer de uma decisão automatizada.

O caso envolvendo a Uber Inc. e a App Drivers & Couriers Union (ADCU), mencionado no Capítulo 3 deste trabalho, é ilustrativo de como a qualificadora da recorribilidade pode ser mobilizada em uma situação prática envolvendo decisões automatizadas. Em julho de 2020, quatro motoristas da Uber associados à ADCU entraram com uma ação em face da empresa na Corte de Amsterdam questionando as decisões automatizadas que resultaram no seu desligamento da plataforma. No caso em questão, os autores chamam a atenção para a forte opacidade em torno da decisão automatizada, uma vez que não foram fornecidas pela empresa explicações sobre o que teria motivado o desligamento ou quais foram os fundamentos da decisão, tendo a Uber se limitado a afirmar que a razão do desligamento estaria relacionada a *atividade fraudulenta*. Por acreditarem que a decisão era injusta e afetava negativamente sua esfera de direitos, os autores decidiram contestá-la com base no direito de revisão contido no art. 22(2) da GDPR, argumentando que as salvaguardas legais não foram observadas e que a empresa falhou em fornecer uma explicação

[650] BRASIL. Superior Tribunal de Justiça. *Recurso Especial nº 1.457.199/RS*. Relator: Min. Paulo de Tarso Sanseverino. Julgado em: 12/11/2014. Data de Publicação: 17/12/2014. Disponível em: https://bdjur.stj.jus.br/jspui/bitstream/2011/114173/REsp1457199.pdf. Acesso em: 8 abr. 2021. p. 80.
[651] BRASIL. Superior Tribunal de Justiça. *Recurso Especial nº 1.457.199/RS*. Relator: Min. Paulo de Tarso Sanseverino. Julgado em: 12/11/2014. Data de Publicação: 17/12/2014. Disponível em: https://bdjur.stj.jus.br/jspui/bitstream/2011/114173/REsp1457199.pdf. Acesso em: 8 abr. 2021.

sobre os critérios da decisão.[652] Assim como no devido processo legal, portanto, a qualificadora da recorribilidade confere ao titular sujeito a uma decisão automatizada a possibilidade de contestá-la, caso acredite que ela incorra em erro ou vício de fundamentação.

A quinta e última qualificadora consiste na garantia de revisão das decisões, estando intimamente ligada à qualificadora da recorribilidade. No devido processo legal, essa garantia pode ser entendida como o direito do indivíduo ter sua decisão revista por outro ente do Estado isento e neutro, diferente daquele que proferiu a decisão objeto de contestação. Nesse sentido, assim como a qualificadora da recorribilidade, a qualificadora da revisão estaria ligada ao princípio do duplo grau de jurisdição. Transposta para a lógica do devido processo informacional, a garantia de revisão pode ser entendida como o direito do titular de obter a revisão da decisão automatizada, para que a decisão que negativamente impactou seus direitos seja revisitada a partir de uma análise neutra e isenta, preferencialmente realizada por uma pessoa natural, e não apenas por outra decisão automatizada. É partindo desse racional, por exemplo, que se originam as críticas à redação atual do texto do art. 20 da LGPD, que prevê o direito à revisão de decisões automatizadas, mas não prevê expressamente que ela deva ser revista por um agente diferente daquele que elaborou a decisão, uma vez que o termo *humana* foi suprimido do texto legal. Nesse ponto, a atual redação do art. 20 da LGPD atende à qualificadora da recorribilidade, ao garantir o direito de contestação das decisões, mas falha em atender à qualificadora da revisão, ao não garantir que a decisão seja revista por um ente diferente daquele que deu origem à decisão com poderes e capacidades suficientes para mudá-la. No contexto da GDPR, a interpretação do art. 22(2) tem sido bastante rigorosa ao entender que é necessária não apenas uma revisão por pessoa natural, mas também ao exigir que a intervenção humana seja significativa, realizada por uma pessoa com poderes para efetivamente mudar os rumos da decisão, e não meramente protocolar, conforme abordamos na seção 3.1.

Este racional foi invocado no citado caso envolvendo a ADCU e Uber para contestar o argumento apresentado pela empresa de que as demissões não

[652] RUSSON, M. Uber sued by drivers over ‘automated robo-firing’. *BBC News*, 26 out. 2020. Disponível em: https://www.bbc.com/news/business-54698858. Acesso em: 1 jun. 2021; e EKKER. *Uber drivers demand access to their personal data*. Disponível em: https://ekker.legal/2020/07/19/uber-drivers-demand-access-to-their-personal-data/. Acesso em: 1 jun. 2021.

resultaram de decisões automatizadas, o que afastaria o regime do direito à revisão previsto na GDPR, uma vez que que havia humanos (*specialized employees*) supervisionando os processos de demissão automatizada. Os autores argumentam, contudo, que a Uber falhou em demonstrar que a intervenção humana exercida teria sido significativa, ou seja, em nível suficiente para descaracterizar a natureza automatizada:

> The ADCU's latest case alleges that Uber kicked drivers off its platform based on algorithms and accused them of "fraudulent activity" without offering any avenue for appeal. *The filing acknowledges Uber's argument that it uses "specialized employees" to assess account deactivations. But the drivers' lawyer, Anton Ekker, argues that Uber has not "further substantiated that this constitutes meaningful human intervention" or that the employees have been trained to understand "how the artificially intelligent system works." The document reads: 'In particular, Uber has not demonstrated that the employees involved in automated decision-making: Have a meaningful influence on the decision, which means, among other things, that they must have the "authority and competence" to oppose this decision*.'[653] (grifo nosso).

Assim como no devido processo legal, portanto, também no devido processo informacional faz-se necessário garantir que a decisão, uma vez contestada, seja reanalisada por um ente distinto daquele que a proferiu, com poderes e capacidades suficientes para reformá-la a partir de uma atuação neutra e isenta.

Juntas, essas cinco qualificadoras — isenção, informação, compreensão, recorribilidade e revisão — somadas ao princípio da prevenção enquanto eixo estruturante, conformam o que aqui denominamos cláusula geral do devido processo informacional, que, aplicado ao contexto de decisões automatizadas, confere aos titulares de dados pessoais e grupos afetados por decisões automatizadas uma série de garantias procedimentais. Conforme salientamos anteriormente, para fins de construção de um modelo de explicabilidade concreto a partir do *framework* proposto, faz-se necessário focar, especialmente, em duas das qualificadoras anteriormente apresentadas, quais sejam, as qualificadoras da informação e compreensão, por entendermos estarem mais diretamente ligadas à construção de uma explicação.

---

[653] SAWERS, Paul. Uber drivers union asks EU court to overrule 'robo firing' by algorithm. *Venture Beat*, 26 out. 2020. Disponível em: https://venturebeat.com/2020/10/26/uber-drivers-union-asks-eu-court-to-overrule-robo-firing-by-algorithm/. Acesso em: 10 jan. 2021.

QUADRO 3 – Qualificadoras do devido processo informacional

PROCESSO	DEVIDO PROCESSO LEGAL	DEVIDO PROCESSO INFORMACIONAL
ISENTO / NEUTRO / IMPARCIAL	Direito de ter sua violação de direitos analisada por um ente imparcial, neutro e isento do Estado, como, por exemplo, um juiz (suspeição e impedimentos, além de outras garantias processuais).	Direito do titular de dados de se sujeitar a um processo automatizado tomador de decisões que seja imparcial, neutro e isento, como, por exemplo, sem vieses discriminatórios.
INFORMADO	Direito de ter acesso às informações sobre o processo, como os autos de um processo judicial, que são acessíveis às partes e, majoritariamente, ao público para cumprir com as obrigações de transparência. Direito de ter acesso aos andamentos do processo para que possa acompanhar o seu andamento.	Direito de ter acesso aos dados que servem de substrato para o processo automatizado, o que pode ser feito por meio práticas de transparência ativa e atendimento à requisições de acesso aos seus dados. Todavia, ter acesso aos dados não necessariamente significa entender e compreender como o processo funciona e estes são utilizados para tomar uma decisão automatizada.
COMPREENSÍVEL / EXPLICADO	Direito de entender as razões, os fundamentos e o racional de uma decisão que lhe afeta, o que é feito, por exemplo, na obrigação do Juiz de expressamente motivar a sua decisão judicial, informando quais provas e argumentos foram aceitos, a valoração destes, e motivos pelos quais os demais não foram aceitos.	Direito à explicação efetiva de como funciona uma decisão automatizada e como esta pode impactar direitos e liberdades fundamentais. Não bastaria apenas o fornecimento de informações, mas compreender o racional, a lógica, a valoração, os pesos e o funcionamentos dos algoritmos. Isso poderia ser atingido de algumas formas, como direitos de acesso mais robustos, análise de códigos fonte, relatórios de impacto algorítmico, linguagem comum, entre outros.
RECORRÍVEL	Direito de recorrer de uma decisão que pode impactar seus direitos, caso não concorde com ela ou ela haja algum tipo de vício nela (duplo grau de jurisdição). Direito de ser ouvido. Há ainda a garantia da inversão do ônus da prova, para que em situações de vulnerabilidade em que o demandante não possui os meios para produzir provas.	Direito de desafiar a decisão automatizada que gerou impacto em seus direitos e liberdades fundamentais. Mas para o desafio ser efetivo, é necessário entender e compreender como a decisão lhe impacta e como ela foi tomada, de forma que somente o controlador é capaz de prover essa fundamentação. Para isso, o direito à explicação.
REVISÁVEL	Direito que a decisão seja revista por outro ente do Estado, também isento e neutro, que também atende aos requisitos anteriores.	Direito de revisão da decisão automatizada, para a decisão que negativamente impactou os seus direitos seja revisitada, também, de forma imparcial, neutra e isenta, preferencialmente por um ente humano, e não apenas por outra decisão automatizada.

Fonte: Elaboração própria

De modo a exemplificar o funcionamento do *framework* ora proposto, nos próximos parágrafos buscaremos colocar em movimento a cláusula geral do devido processo informacional, concentrando-nos nas qualificadoras da informação e da compreensão, tomando como objeto um dos domínios de aplicação de decisões automatizadas anteriormente apresentado, o de segurança pública e persecução penal, cuja lógica pode ser replicada para os demais domínios apresentados. Nosso objetivo é exemplificar como o *framework* aqui proposto pode ser útil à construção de um modelo concreto de explicabilidade à luz de diferentes contextos de aplicação, diferentes níveis de risco e opacidade e diferentes perfis de destinatários de uma explicação.

Considerado o domínio de aplicação a ser analisado (segurança pública e persecução penal), poderíamos indagar, para utilizar um exemplo aqui citado, como conceber um modelo de explicabilidade útil no contexto de implementação do COMPAS, o *software* de análise de risco de reincidência criminal utilizado pela justiça norte-americana?

Comecemos analisando, primeiramente, o que seria um modelo de explicabilidade adequado em função dos destinatários da explicação. Conforme

destacado anteriormente, propomos uma análise do destinatário da explicação baseada em quatro fatores: (i) contexto da interação, (ii) necessidades, (iii) complexidade do sistema e (iv) possibilidade de desafiar a decisão.

Quando tratamos do domínio da segurança pública e persecução penal, faz-se necessário ter em mente que, em razão do próprio contexto da interação, não há apenas um único destinatário das obrigações de transparência, mas múltiplos destinatários, que, para fins da análise aqui empreendida, podem ser definidos em duas categorias: a sociedade e a opinião pública em geral e o cidadão diretamente sujeito a uma decisão automatizada.

Conforme salientamos na seção 5.2.1.3, quando analisamos o contexto da interação estamos a tratar, especialmente, do domínio de aplicação considerado, dos riscos e impactos que lhes são inerentes e dos direitos envolvidos. No contexto de decisões automatizadas aplicadas para fins de segurança pública e persecução penal, estamos a tratar do próprio exercício do poder de polícia do Estado, que, por meio de novas tecnologias, limita o exercício de direitos e garantias fundamentais e das liberdades individuais dos cidadãos.

Ao fazer uso de decisões automatizadas no campo da segurança pública e persecução penal, o Estado coloca em xeque não apenas os direitos e garantias de um único cidadão, mas valores, bens jurídicos e garantias socialmente tutelados, que têm como titulares toda uma coletividade difusa e indeterminada. Nesse contexto, um primeiro destinatário da explicação seria um destinatário difuso, não determinado, abarcando a sociedade como um todo, a quem o Estado deve justificar, fundamentar e prestar contas a respeito do exercício do seu poder de polícia.

Ao lado deste destinatário amplo, haveria ainda um destinatário individualizado, entendido como aquele titular/cidadão diretamente sujeito a uma decisão automatizada para fins de cálculo de risco de reincidência e cálculo de pena, por exemplo. Como vimos no item A, pelo alto risco de discriminação envolvido, bem como pelo alto potencial de restrição a direitos fundamentais, liberdades individuais e garantias processuais, o emprego de decisões automatizadas nesse domínio deve ser visto como de altíssimo risco. Em termos práticos, isso exige não apenas cautela quanto ao seu emprego, como também o dever de adoção de salvaguardas e obrigações de transparência e prestação de contas mais robustas e em maior número.

Um segundo fator a ser considerado são as necessidades de cada destinatário da explicação. Obviamente, o tipo de explicação que se requer para fins de controle

social e escrutínio público sobre o exercício do poder de polícia do Estado será bastante diferente, tanto em forma quanto em conteúdo, do tipo de explicação que um indivíduo diretamente afetado necessita receber para compreender e contestar determinada decisão.

Quando considerado um destinatário difuso, ganham relevo, sobretudo, mas não unicamente, as obrigações de transparência mais estruturais e de natureza *ex ante*, que permitem um controle social e alguma previsibilidade sobre os impactos da atuação do Estado. Entram em jogo, neste contexto, a adoção de práticas de *Privacy by Design* e *Privacy by Default* no desenho de ferramentas e de processos voltados à segurança pública e persecução penal, a elaboração de Relatórios de Impacto à Proteção de Dados Pessoais prévios, nos termos do art. 4º, § 3º, da LGPD[654], a elaboração de auditorias para verificação de aspectos discriminatórios nas soluções tecnológicas empregadas e a documentação, por meio de trilhas de auditoria algorítmica, de todos os passos seguidos pelo sistema na elaboração da decisão automatizada, até mesmo como garantia do contraditório e da ampla defesa e do princípio da motivação das decisões judiciais.

Quando tratamos do indivíduo diretamente afetado por uma decisão automatizada, ganham relevo, sem prejuízo da observância das garantias *ex ante*, sobretudo as ferramentas de proteção *ex post* e as explicações locais, voltadas a explicitar os fundamentos e o racional das decisões, embora explicações globais, acerca do funcionamento do algoritmo, também se façam necessárias em alguma medida. Neste contexto também ganham relevo as requisições relativas aos direitos de confirmação da existência de tratamento e de acesso aos dados pessoais, bem como a requisição de realização de auditorias para verificação de aspectos discriminatórios. A determinação do sujeito destinatário da explicação, portanto, é de fundamental importância dado que, em última análise, dela decorrem as necessidades que precisarão ser atendidas por cada tipo de explicação.

Em terceiro lugar, deve ainda ser considerada a complexidade do sistema, pois,

---

654 "Art. 4º. [...] §3º. Esta Lei não se aplica ao tratamento de dados pessoais: [...] III — realizado para fins exclusivos de: a) segurança pública; b) defesa nacional; c) segurança do Estado; ou d) atividades de investigação e repressão de infrações penais;[...]§ 3º A autoridade nacional emitirá opiniões técnicas ou recomendações referentes às exceções previstas no inciso III do caput deste artigo e deverá solicitar aos responsáveis relatórios de impacto à proteção de dados pessoais." (BRASIL. *Lei nº 13.709, de 14 de agosto de 2018*. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF: Presidência da República, 14 ago. 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm. Acesso em: 07. nov. 2020).

em última análise, o grau de opacidade existente traz implicações quanto a quem deverá ser direcionada a explicação, qual tipo de explicação deverá ser elaborada ou até mesmo se será possível elaborar algum tipo de explicação. Como vimos no item A, frequentemente, as tecnologias empregadas nas aplicações de reconhecimento facial encerram um alto grau de opacidade, geralmente caracterizada como modelos não supervisionados de aprendizado de máquina, no qual é possível conhecer os dados de entrada e os resultados (*outputs*) do sistema, mas cujos processos intermediários realizados pela máquina não são compreensíveis ou exprimíveis de forma lógica. Como vimos no Capítulo 3, sistemas desse tipo exigem um alto grau de especialização técnica para serem compreendidos, e, por vezes, nem mesmo seus desenvolvedores possuem plena consciência de como os *outputs* são gerados. Como pontuado no Capítulo 4, a opacidade inerente dos sistemas de reconhecimento facial costuma ser apontada por críticos como uma das razões para o seu completo banimento em domínios de alto risco[655], como em aplicações voltadas para segurança pública e persecução penal. Quanto maior a complexidade do sistema, portanto, maiores devem ser as obrigações de transparência e maior a especialização necessária para compreendê-lo (ou maior o esforço para traduzir as explicações para o público amplo e para o usuário comum).

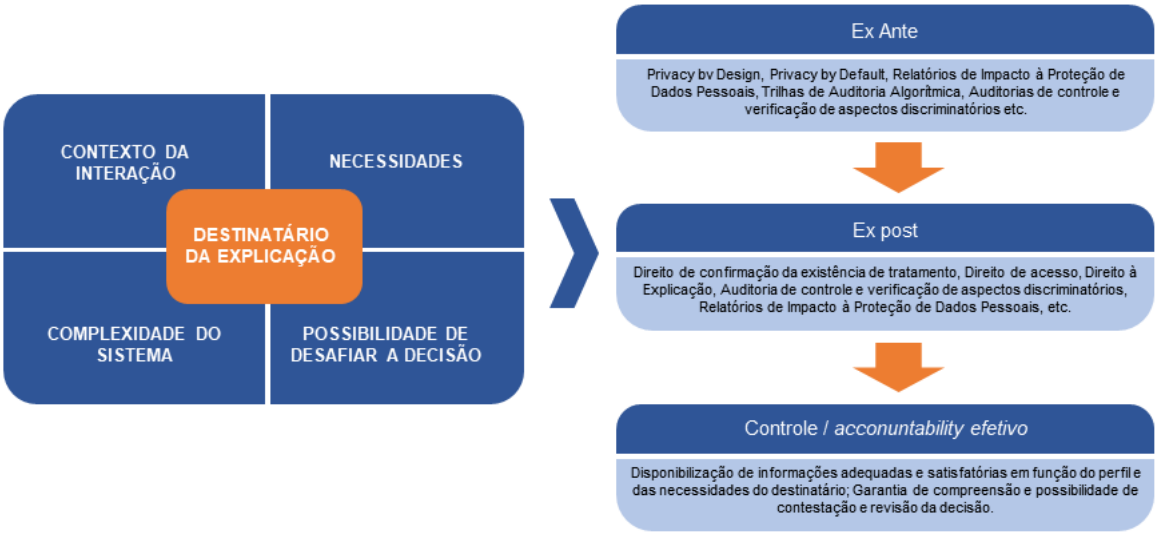
Um quarto e último aspecto do destinatário da explicação está em analisar se, com base no contexto da interação, nas necessidades a serem supridas e no grau de opacidade do sistema, ele estaria apto a desafiar a decisão. Este fator deve estar no horizonte de qualquer modelo de explicabilidade, uma vez que sua função deve ser, em última análise, munir o destinatário de subsídios para compreender e, a partir desta compreensão, avaliar a pertinência de contestar e solicitar a revisão de determinada decisão automatizada.

No quadro abaixo, buscamos sistematizar em um fluxograma o racional aqui apresentado, que, juntamente com as considerações preliminares apresentadas no início deste tópico, poderá guiar a construção de modelos de explicabilidade em função de diferentes destinatários, diferentes domínios de aplicação e diferentes graus de risco de atividade e de opacidade algorítmica.

---

[655] RUDIN, C. Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead. *ArXiv [cs, stat]*, [*S. l.*], 2019. Disponível em: http://arxiv.org/abs/1811.10154. Acesso em: 10 ago. 2020.

FIGURA 1 – Garantia das qualidades da informações e da compreensão em função
Dos destinatários da explicação em perspectiva ex ante e ex post



Fonte: Elaboração própria

Na primeira dimensão do fluxograma, à esquerda, encontra-se situado o destinatário da explicação, aqui tomado como centro e ponto de partida do modelo de explicabilidade descrito, bem como os quatro fatores a ele relacionados (contexto da interação, necessidades, complexidade do sistema e possibilidade de desafiar a decisão). Avançando em relação a essa dimensão, nos deparamos com dois momentos nos quais as qualificadoras da informação e da compreensão deverão ser operacionalizadas, levando em conta os diferentes instrumentos técnico-jurídicos à disposição na "caixa de ferramentas". Em última análise, a aplicação dessas ferramentas, seja em uma perspectiva *ex ante*, seja numa perspectiva *ex post*, deve levar a resultados mensuráveis, quais sejam, a garantia da autodeterminação informativa e de uma *accountability* efetiva, garantindo ao titular ou grupo afetado a oportunidade de compreender e desafiar determinada decisão automatizada.

5.3 SÍNTESE DO CAPÍTULO

Não obstante as diversas limitações apresentadas no Capítulo 4, o presente capítulo buscou articular formas de implementação e garantia do direito à explicação

a partir do direito vigente. Para tanto, na seção 5.1, apresentamos os diferentes instrumentos existentes nas legislações de proteção de dados pessoais, com especial enfoque no tratamento dado a cada instrumento na LGPD, que podem ser empregados para dar concretude ao direito à explicação, sendo eles: os diferentes papéis e obrigações atribuídos aos agentes de tratamento, os direitos morais dos titulares de dados e as obrigações de transparência, o papel da *accountability* e da responsabilidade demonstrável, os relatórios de impacto à proteção de dados pessoais, a regulaç*ão ex ante* pelo código na forma de uma "*explainability by design*" e as diferentes modalidades de auditorias algorítmicas. Em seguida, na seção 5.2, buscamos articular essas diferentes ferramentas técnico-jurídicas apresentadas na seção 5.1, bem como o debate teórico sobre o devido processo informacional, apresentado sobretudo no Capítulo 2, para propor um *framework* de explicabilidade. Antes da apresentação do *framework* em si, discorremos sobre algumas das diretrizes e pressupostos que, a nosso ver, devem ser considerados na concepção de um *framework* de explicabilidade. Neste sentido, na seção 5.2.1, pontuamos: (i) que o *framework* de explicabilidade proposto não pretende ser uma solução única e genérica, mas uma "caixa de ferramentas", composta por diferentes instrumentos técnico-jurídicos, que podem ser combinados para a construção de um modelo de explicação em uma determinada situação concreta; (ii) que o *framework* proposto deve partir de uma abordagem contextual que considere risco e opacidade do sistema como variáveis norteadoras para a construção de um modelo de explicabilidade; e (iii) que a instrumentalização do direito à explicação deve partir de uma abordagem centrada no destinatário da explicação, considerando-se o contexto da interação, as necessidades do destinatário, a complexidade do sistema automatizado em questão e a possibilidade do sujeito desafiar a decisão. Partindo dessas considerações e do debate teórico apresentado no Capítulo 2, buscou-se desenvolver uma proposta de *framework* de explicabilidade a partir da cláusula geral do devido processo informacional. Para tanto, foram inicialmente delimitados e explicitados os elementos constitutivos da cláusula geral do devido processo informacional. Em seguida, esses elementos foram contextualizados e colocados em perspectiva com os elementos constitutivos da cláusula geral do devido processo legal, a partir de uma literatura de matriz constitucional. Partimos de uma concepção de cláusula geral do devido processo informacional constituída pelos seguintes elementos: (i) isenção, (ii) informação, (iii) compreensão, (iv) recorribilidade e (v) revisão. Após apresentados e

contextualizados cada um dos elementos constitutivos, o trabalho optou, conforme explicitado, por aprofundar em apenas dois deles, quais sejam, a informação e a compreensão, por entendermos que estarem mais intimamente relacionados à construção de uma explicação. A partir desse recorte, cada um dos elementos foi então contextualizado/colocado em prática a partir de um dos exemplos apresentados na seção 5.1 (domínios de aplicação A, B e C). Por fim, o trabalho buscou explorar como a informação e a compreensão poderiam ser garantidos a partir de uma perspectiva tanto *ex ante* quanto *ex post*, considerando-se os diferentes instrumentos técnico-jurídicos apresentados na seção 5.1.

CONCLUSÃO

A presente tese visou explorar a hipótese da existência no ordenamento jurídico brasileiro de um direito à explicação no contexto de decisões automatizadas que tenham como substrato o uso de dados pessoais. Este pode ser entendido como o direito a receber informações suficientes e inteligíveis que permitam ao titular dos dados e à sociedade entenderem e compreenderem a lógica, a forma e os critérios utilizados para tratar dados pessoais em decisões automatizadas e entender e prever os seus impactos, com o fim de evitar práticas discriminatórias, inadequadas, ilegítimas e indesejadas, que podem ter impacto no plano individual e coletivo em direitos e liberdades fundamentais.

Todavia, como explorado, não há consenso sobre a real existência de tal direito, uma vez que ele não estaria positivado expressamente na LGPD. Por isso buscamos explorar a hipótese de que, sim, é possível reconhecer a existência de tal direito em nosso ordenamento, partindo de uma leitura tanto da nossa Lei Geral de Proteção de Dados, de um regime de *accountability,* de obrigações de transparência e da cláusula geral do devido processo informacional. Para tanto, no Capítulo 2, discutimos o que é o direito à explicação no cenário brasileiro e europeu, a partir de fontes positivadas, como a LGPD e a GDPR. Neste sentido, o capítulo buscou desenvolver o argumento de que é possível defender a existência desse direito no ordenamento brasileiro a partir de uma leitura sistemática da LGPD, das legislações setoriais de proteção de dados pessoais, da Constituição e da jurisprudência dos tribunais superiores. No âmbito da UE, discutimos de que forma ele pode ser extraído da leitura do texto da GDPR e a interpretação que tem sido dada a esse direito pelas autoridades de *enforcement*. Além das fontes legais, o capítulo buscou apresentar ainda outras formas de reconhecimento de um direito à explicação, como o conceito de autodeterminação informativa, a cláusula geral dos direitos da personalidade e do devido processo informacional, que também podem servir como fundamentos úteis ao reconhecimento de um direito à explicação.

No cenário brasileiro, especificamente, demonstramos ser possível reconhecer a existência de um direito à explicação a partir de uma leitura sistemática da LGPD e das legislações setoriais nacionais de proteção de dados, que, em conjunto, oferecem um robusto leque de direitos, princípios e ferramentas que, uma vez articulados e mobilizados, permitem garantir e instrumentalizar um direito à explicação. Por vezes,

sobretudo em razão do maior escopo do conceito de dado pessoal positivado na legislação brasileira, que, pode-se defender, adotou a abordagem consequencialista, o direito à explicação na LGPD assume contornos ainda mais amplos que aqueles previstos na GDPR. Todavia, quando comparados outros aspectos, como a ausência de uma previsão expressa à revisão de decisões automatizadas por pessoa natural, o alto relevo dado ao segredo de negócio e a abertura à discricionaridade da ANPD para a realização de auditorias algorítmicas criada pela redação do art. 20, tem-se um direito à explicação na LGPD que pode sofrer mais restrições do que em relação àquele contido na GDPR.

Reconhecida a existência do direito à explicação, é necessário ainda identificar as limitações que podem existir para a sua efetiva aplicação, tais como as impostas pela legislação de propriedade intelectual, principalmente segredo de negócio, o próprio conceito de dado pessoal, a opacidade inerente dos sistemas complexos de inteligência artificial e *machine learning,* a capacidade de compreensão dos titulares de dados das eventuais explicações concedidas e as restrições institucionais das entidades supervisoras para analisarem os sistemas baseados em decisões automatizadas. No capítulo 4, buscamos analisar detidamente cada uma dessas limitações e suas consequências para o reconhecimento, aplicação e instrumentalização de um efetivo direito à explicação. Quanto às limitações legais associados ao segredo de negócio, buscamos argumentar que a proteção – legítima – aos direitos de propriedade intelectual não deve ser absoluta e não poderia servir de escudo para obstaculizar o cumprimento a obrigações de transparência e a garantia do próprio direito à explicação. Os direitos de propriedade intelectual limitam em alguma medida o direito à explicação, mas devem ser sempre sopesados em relação aos outros interesses e direitos fundamentais presentes em determinado caso concreto. Em última análise, buscamos oferecer ainda soluções conciliatórias para o conflito entre esses dois polos de interesses, consistente na realização de auditorias algorítmicas por terceiros e entes neutros, na disponibilização de diferentes versões (públicas e de acesso restrito) de relatórios de impacto e auditorias algorítmicas, dentre outros. Argumentamos, ademais, pela necessidade de adoção de um conceito mais robusto de dado pessoal, menos focado na ‘identificabilidade’ de determinada pessoa natural, e mais preocupado em mitigar os efeitos que determinado tratamento de dados, sejam eles pessoais ou não, possam ter sobre a esfera de direitos de indivíduos e grupos. Quanto aos limites de compreensão do titular e a complexidade

inerente a sistemas algorítmicos, argumentamos que o ônus de explicabilidade deve ser maior quanto maiores forem os níveis de opacidade do sistema. Somado a isso, é necessário que as explicações sejam construídas tendo como referência o titular, uma persona média, considerados seu nível de conhecimento, suas limitações, seus objetivos e suas necessidades, de modo a garantir a construção de uma explicação de natureza instrumental e efetiva do ponto de vista prático, que possibilite ao titular a oportunidade de entender os impactos que determinada decisão automatizada pode ter em seus direitos e solicitar a revisão dessa decisão, se assim o desejar. Analisamos, ainda, as diversas limitações institucionais para o *enforcement* do direito à explicação: uma vez reconhecido, é necessário que este direito seja efetivamente garantido, e tanto os recursos humanos e financeiros de uma autoridade de supervisão, quanto o desenho institucional e os atributos de independência e autonomia acabam influenciando na qualidade do *enforcement* e no grau de efetivação desse direito.

Não obstante tais limitações, analisamos de que forma diferentes instrumentos existentes nas legislações de proteção de dados pessoais poderiam ser empregados para dar concretude ao direito à explicação, com foco numa regula*ção ex ante* na forma de uma "*explainability by design*" e as diferentes modalidades de auditorias algorítmicas e relatórios de impacto, conforme apresentamos na primeira parte do capítulo 5 (seção 5.1).

Por fim, confirmada a hipótese da existência do direito à explicação, buscou-se desenvolver um framework de explicabilidade a partir da cláusula geral do devido processo informacional, na forma de uma "caixa de ferramentas", composta por diferentes instrumentos técnico-jurídicos, que podem ser combinados para a construção de um modelo de explicação em uma determinada situação concreta.

Referido *framework* é apenas um ponto de partida para a construção de instrumentos mais robustos que possam colaborar para identificar os impactos que decisões automatizadas podem ter em nossos direitos e liberdades fundamentais numa sociedade intrinsicamente influenciada por algoritmos. O desenvolvimento desses instrumentos, todavia, depende fortemente do reconhecimento que as atuais leis de proteção de dados, incluindo a LGPD e a GDPR, podem ser insuficientes para endereçar tais problemas, que inclusive podem estar fora dos seus escopos. Por isso, um grande tema a ser melhor explorado em pesquisas futuras é o próprio conceito de dados pessoais que serve de substrato para todas as normas de proteção de dados.

REFERÊNCIAS BIBLIOGRÁFICAS

ACCESS NOW; AMNESTY INTERNATIONAL. *The Toronto Declaration:* protecting the right to equality and non-discrimination in machine learning systems. Toronto, 2018. Disponível em: https://www.accessnow.org/cms/assets/uploads/2018/08/The-Toronto-Declaration_ENG_08-2018.pdf. Acesso em: 15 dez. 2020.

ACDU. Uber Drivers Take Unprecedented International Legal Action To Demand Their Data. *ACDU*, 20 jul. 2020. Disponível em: https://www.adcu.org.uk/news-posts/uber-drivers-take-unprecedented-international-legal-action-to-demand-their-data. Acesso em: 31 jul. 2020.

ACKERMAN, J. M.; SANDOVAL-BALLESTEROS, I. E. The Global Explosion of Freedom of Information Laws. *Administrative Law Review*, [*S. l.*], v. 58, n. 1, p. 85–130, 2006.

ADA INSTITUTE. *Examining Tools for assessing algorithmic systems the Black Box*. [*S. l.*]: [*s. n.*], 2020.

ADADI, A.; BERRADA, M. Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *IEEE Access*, [*S. l.*], v. 6, p. 52138–52160, 2018. Disponível em: https://doi.org/10.1109/ACCESS.2018.2870052. Acesso em: 21 jun. 2021.

AEPD. Guía práctica para las evaluaciones de Impacto en la Protección de los datos sujetas al RGPD. *AEPD*, 6 maio 2021. Disponível em: https://www.aepd.es/es/derechos-y-deberes/cumple-tus-deberes/medidas-de-cumplimiento/evaluaciones-de-impacto. Acesso em: 6 jun. 2021.

AGENCIA DE ACCESO A LA INFORMACIÓN PÚBLICA; UNIDAD REGULADORA Y DE CONTROL DE DATOS PERSONALES. *Guía de Evaluación de Impacto en la Protección de Datos*. [*S. l.*]: Agencia de acceso a la información pública; Unidad reguladora y de control de Datos Personales, 2020. Disponível em: https://www.argentina.gob.ar/sites/default/files/guia_final.pdf. Acesso em: 9 jun. 2021.

AHMED, N.; WAHED, M. The De-democratization of AI: Deep Learning and the Compute Divide in Artificial Intelligence Research. *ArXiv*, 2020. Disponível em: http://arxiv.org/abs/2010.15581. Acesso em: 9 nov. 2020.

ALEXY, R. *Teoria dos direitos fundamentais*. São Paulo: Malheiros, 2008.

ALHADEFF, J.; VAN ALSENOY, B.; DUMORTIER, J. The Accountability Principle in Data Protection Regulation: Origin, Development and Future Directions. *In*: GUAGNIN, D. *et al.* (org.). *Managing Privacy through Accountability*. London: Palgrave Macmillan UK, 2012. p. 49–82. Disponível em: https://doi.org/10.1057/9781137032225_4. Acesso em: 16 dez. 2020.

ALIMONTI, V. Autodeterminação informacional na LGPD: antecedentes, influências e desafios. *In*: DONEDA, D.; MENDES, L. S.; CUEVA, R. V. B. (Coords.). *Lei Geral de Proteção de Dados (Lei nº 13.709/2018) — A caminho da efetividade:* contribuições para a implementação da LGPD. São Paulo: Thomsom Reuters, 2020.

ANANNY, M.; CRAWFORD, K. Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, v. 20, n. 3, p. 973–989, 2016.

ANGWIN, J. *et al*. Machine Bias. *ProPublica*, 23 maio 2016. Disponível em: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. Acesso em: 27 maio. 2020.

ARRUDA, C. S. L. de. Breve estudo hermenêutico-epistemológico da cláusula do "devido processo legal". *Revista CEJ*, Brasília, Ano XXI, n. 73, set./dez. 2017.

ARTICLE 29 DATA PROTECTION WORKING PARTY*. Opinion 04/2014 on surveillance of electronic communications for intelligence and national security purposes.* Bruxelas: European Commission, 10 abr. 2014. Disponível em: https://archiwum.giodo.gov.pl/pl/file/5982. Acesso em: 27 jul. 2020.

ASGHARI, H.; VAN EETEN, M.; MAHIEU, R. *Collectively Exercising the Right of Access*: Individual Effort, Societal Effect. Rochester, NY: Social Science Research Network, 2017. SSRN Scholarly Paper. Disponível em: https://doi.org/10.2139/ssrn.3107292. Acesso em: 9 dez. 2020.

AUTOMÁTICO. *In*: MICHAELIS. São Paulo: Editora Melhoramentos, 2021. Disponível em: https://michaelis.uol.com.br/busca?id=bK57. Acesso em: 14 jun. 2021.

AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS. ANPD e Senacon assinam acordo de cooperação técnica. *ANPD*, 22 mar. 2021. Disponível em: https://www.gov.br/anpd/pt-br/assuntos/noticias/anpd-e-senacon-assinam-acordo-de-cooperacao-tecnica. Acesso em: 16 maio 2021.

AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS. Cade, MPF, ANDP e Senacon recomendam que WhatsApp adie entrada em vigor da nova política de privacidade. *ANPD*, 7 maio 2021. Disponível em: https://www.gov.br/anpd/pt-br/assuntos/noticias/cade-mpf-anpd-e-senacon-recomendam-que-whatsapp-adie-entrada-em-vigor-da-nova-politica-de-privacidade. Acesso em: 21 jun. 2021.

AUXÍLIO emergencial: Bases de dados utilizadas pela Dataprev. *Dataprev*. 20 jun. 2020. Disponível: https://portal2.dataprev.gov.br/bases-de-dados-ultilizadas-no-processamento-do-auxilio-emergencial. Acesso em: 23 nov. 2020.

AZEVEDO, I. T. R.; SILVA, T. A. da. Reflexões sobre tomada de decisão e livre arbítrio sob a ótica da neurociência e seus efeitos no sistema punitivo. *LINKSCIENCEPLACE* - Interdisciplinary Scientific Journal, [*S. l.*], v. 1, n. 1, 2014. Disponível em: http://revista.srvroot.com/linkscienceplace/index.php/linkscienceplace/article/view/16. Acesso em: 14 jun. 2021.

BAKER, J. J. Beyond the Information Age: The Duty of Technology Competence in the Algorithmic Society. *South Carolina Law Review*, v. 69, n. 557, 2018. Disponível em: https://ttu-ir.tdl.org/handle/2346/73913. Acesso em: 21 out. 2020.

BALL, C. What Is Transparency? *Public Integrity*, [*S. l.*], v. 11, n. 4, p. 293–308, 2009. Disponível em: https://doi.org/10.2753/PIN1099-9922110400. Acesso em: 21 jun. 2021.

BENNET, C. *Regulating privacy*: data protection and public policy in Europe and the United States. Ithaca, New York: Cornell University Press, 1992.

BENNET, C. The Accountability Approach to Privacy and Data Protection: Assumptions and Caveats. *In*: GUAGNIN, D. *et al.* (org.). *Managing Privacy through Accountability*. London: Palgrave Macmillan UK, 2012. p. 33–48. E-book.

BINNS, R. Data Protection Impact Assessments: A Meta-Regulatory Approach. *International Data Privacy Law*, v. 7, n. 1, p. 22-35, 2017. Disponível em: https://papers.ssrn.com/abstract=2964242. Acesso em: 24 abr. 2021.

BIONI, B. Xeque-Mate: o tripé de proteção de dados pessoais no xadrez das iniciativas legislativas no Brasil. *GPOPAI/USP*, [*S. l.*], 2015. Disponível em: https://www.researchgate.net/publication/328266374_Xeque-Mate_o_tripe_de_protecao_de_dados_pessoais_no_xadrez_das_iniciativas_legislativas_no_Brasil. Acesso em: 27 ago. 2020.

BIONI, B. R. Compreendendo o conceito de anonimização e dado anonimizado. *Revista do Advogado*, n. 144, nov. 2019. Disponível em: https://aplicacao.aasp.org.br/aasp/servicos/revista_advogado/paginaveis/144/20/index.html. Acesso em: 21 jul. 2021.

BIONI, B. R. *Proteção de dados pessoais:* a função e os limites do consentimento. 1. ed. Rio de Janeiro: Forense, 2019.

BIONI, B. R. *Proteção de dados pessoais*: a função e os limites do consentimento. 2. ed. Rio de Janeiro: Forense, 2021.

BIONI, B.; MARTINS, P. Devido processo informacional: um salto teórico-dogmático necessário? *Jota*, 15 jul. 2020. Disponível em: https://www.jota.info/opiniao-e-analise/artigos/devido-processo-informacional-um-salto-teorico-dogmatico-necessario-15072020. Acesso em: 15 jul. 2020.

BIONI, B.; MENDES, L. S. Regulamento Europeu de Proteção de Dados Pessoais e a Lei Geral brasileira de Proteção de Dados: mapeando convergências na direção de um nível de equivalência. *In*: TEPEDINO, G.; FRAZÃO, A.; OLIVA, M. D. (Coords.). *Lei Geral de Proteção de Dados Pessoais e suas repercussões no Direito Brasileiro*. 2. ed. São Paulo: Thomson Reuters, 2020.

BOBBIO, N. Sobre os fundamentos dos direitos do homem. *In:* BOBBIO, N. *A Era dos Direitos*. 7. ed. Apresentação: Celso Lafer; Tradução: Carlos Nelson Coutinho. Rio de Janeiro: Elsevier, 2004.

BOBBIO, N. *Teoria do ordenamento jurídico*. Apresentação: Tércio Sampaio Ferraz Júnior; Tradução: Maria Celeste C. J. Santos; Rev. téc.: Cláudio de Cicco. 6. ed. Brasília: Editora Universidade de Brasília, 1995.

BOHLENDER, D.; KÖHL, M. A. Towards a Characterization of Explainable Systems. *ArXiv,* 2019. Disponível em: http://arxiv.org/abs/1902.03096. Acesso em: 6 out. 2020.

BORGES, J. L. Sobre o Rigor na Ciência. *In*: BORGES, J. L. *História Universal da Infâmia*. Lisboa: Assírio e Alvim, 1982.

BOUCHER, P. “Safari” ou la chasse aux Français. *Le Monde*, p. 9, 21 mar. 1974. Disponível em: https://www.cnil.fr/sites/default/files/atoms/files/le_monde_0.pdf. Acesso em: 21 jun. 2021.

BRASIL. *Lei nº 8.078, de 11 de setembro de 1990*. Dispõe sobre a proteção do consumidor e dá outras providências. Brasília, DF: Presidência da República, 12 set. 1990. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l8078compilado.htm. Acesso em: 8 abr. 2021.

BRASIL. *Lei nº 9.279, de 14 de maio de 1976*. Regula direitos e obrigações relativos à propriedade industrial. Brasília, DF: Presidência da República, 14 maio 1996. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l9279.htm. Acesso em: 26 abr. 2021.

BRASIL. *Lei nº 12.414, de 09 de junho de 2011*. Disciplina a formação e consulta a bancos de dados com informações de adimplemento, de pessoas naturais ou de pessoas jurídicas, para formação do histórico de crédito. Brasília, DF: Presidência da República, 10 jun. 2011. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12414.htm. Acesso em: 17 ago. 2020.

BRASIL. Superior Tribunal de Justiça. *Recurso Especial nº 22.337 – RS*. Relator: Ruy Rosado Aguiar. Data de Publicação: 20/03/1995. *In: R. Sup. Trib. Just. Brasília*, a.8 (77), jan. 1996.

BRASIL*. Lei nº 13.709, de 14 de agosto de 2018*. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF: Presidência da República, 14 ago. 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 07. nov. 2020.

BRASIL. *Lei Complementar nº 166, de 8 de abril de 2019*. Altera a Lei Complementar nº 105, de 10 de janeiro de 2001, e a Lei nº 12.414, de 9 de junho de 2011, para dispor sobre os cadastros positivos de crédito e regular a responsabilidade civil dos operadores. Brasília, DF: Presidência da República, 08 abr. 2019. Disponível em: http://www.planalto.gov.br/ccivil_03/LEIS/LCP/Lcp166.htm#art2. Acesso em: 8 abr. 2021.

BRASIL. *Contrato Administrativo nº 12/2020*. Ministério da Cidadania, Secretaria Executiva, Subsecretaria de Assuntos Administrativos. Processo Administrativo nº 71000.022387/2020-55. Diário Oficial da União, Brasília-DF, Seção 3, nº 91, de 14 de maio de 2002, p. 5. Disponível em: http://www.mds.gov.br/webarquivos/acesso_informacao/contratos/2020/12.2020/Contrato%20Adminisntrativo%20n%C2%BA%2012.2020%20-%20DATAPREV.pdf.

Acesso em: 23 nov. 2020.

BRASIL. *Decreto nº 10.316, de 7 de abril de 2020*. Brasília: Presidência da República, 07 abr. 2020. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2019-2022/2020/decreto/d10316.htm. Acesso em: 24 fev. 2021.

BRASIL. *Portaria nº 351, de 7 de abril de 2020*. Ministério da Cidadania. Diário Oficial da União, Brasília-DF, Seção 1 — Extra, edição 67-B, publicado em 7 abr. 2020, p. 13. Disponível em: https://www.in.gov.br/en/web/dou/-/portaria-n-351-de-7-de-abril-de-2020-251562808. Acesso em: 24 fev. 2021.

BRASIL. Autoridade Nacional de Proteção de Dados. *Portaria nº 11, de 27 de janeiro de 2021*. Torna pública a agenda regulatória para o biênio 2021-2022. Diário Oficial da União, Brasília-DF, Seção 1, edição 19, publicado em 28 jan. 2021, p. 3. Disponível em: https://www.in.gov.br/en/web/dou/-/portaria-n-11-de-27-de-janeiro-de-2021-301143313. Acesso em: 21 maio 2021.

BRASIL. Câmara dos Deputados. Sancionada, com nove vetos, lei que cria Autoridade Nacional de Proteção de Dados. *Câmara dos Deputados*, 9 jul. 2019. Disponível em: https://www.camara.leg.br/noticias/561908-SANCIONADA,-COM-NOVE-VETOS,-LEI-QUE-CRIA-AUTORIDADE-NACIONAL-DE-PROTECAO-DE-DADOS. Acesso em: 28 abr. 2021.

BRASIL. Ministério da Justiça. *Portaria nº 5 de 27 de agosto de 2002*. Dispõe sobre cláusulas abusivas em contratos de vendas de produtos e prestação de serviços. Diário Oficial da República Federativa do Brasil, Brasília-DF, 28 ago. 2002. Disponível em: https://www.procon.go.gov.br/legislacao/portarias/portaria-n%C2%BA-5-27-08-2002-mj-sde-clausulas-abusivas-nome-de-consumidor-a-banco-dedados.html. Acesso em: 20 out. 2018

BRASIL. Superior Tribunal de Justiça. *Recurso Especial nº 1.200.105 — AM 2010/0111335-0*. Relator: Min. Paulo de Tarso Sanseverino. Data de Julgamento: 21/05/2013. Data de Publicação: 27/05/2013. Disponível em: https://stj.jusbrasil.com.br/jurisprudencia/23342165/embargos-de-declaracao-no-recurso-especial-edcl-no-resp-1200105-am-2010-0111335-0-stj/inteiro-teor-23342166?ref=amp. Acesso em: 17 ago. 2020.

BRASIL. Superior Tribunal de Justiça. *Recurso Especial nº 1.419.697/RS*. Relator: Min. Paulo de Tarso Sanseverino. Julgado em: 12/11/2014. Data de Publicação: 17/11/2014. Disponível em: https://bdjur.stj.jus.br/jspui/bitstream/2011/114173/REsp1419697.pdf. Acesso em: 8 abr. 2021.

BRASIL. Superior Tribunal de Justiça. *Recurso Especial nº 1.457.199/RS*. Relator: Min. Paulo de Tarso Sanseverino. Julgado em: 12/11/2014. Data de Publicação: 17/12/2014. Disponível em: https://bdjur.stj.jus.br/jspui/bitstream/2011/114173/REsp1457199.pdf. Acesso em: 8 abr. 2021.

BRASIL. Superior Tribunal de Justiça. *Recurso Especial nº 1.304.736 RS*

2012/0031839-3. Relator: Min. Luis Felipe Salomão. Data de Publicação: 30/03/2015. Disponível em: https://stj.jusbrasil.com.br/jurisprudencia/178798658/recurso-especial-resp-1304736-rs-2012-0031839-3. Acesso em: 17 ago. 2020.

BRASIL. Superior Tribunal de Justiça. *Súmula 550*. Julgado em: 14/10/2015. Data de Publicação: 19/10/2015.

BRASIL. Superior Tribunal de Justiça. *Recurso Especial nº 1.304.736/RS*. Relator: Min. Luís Felipe Salomão. Julgado em: 24/02/2016. Data de Publicação: 20/03/2016. Disponível em: https://processo.stj.jus.br/processo/revista/documento/mediado/?componente=ITA&sequencial=1489563&num_registro=201200318393&data=20160330&peticao_numero=-1&formato=PDF. Acesso em: 8 abr. 2021.

BRASIL. Superior Tribunal de Justiça. *Recurso em Mandado de Segurança nº 61.306 – RJ (2019/0199274-6)*. Relator: Min. Luís Felipe Salomão. Julgado em: 4/12/2019.

BRASIL. Supremo Tribunal Federal. *Medida Cautelar na Ação Direta de Inconstitucionalidade nº 6.387 Distrito Federal*. Relator: Min. Rosa Weber. Data de Julgamento: 24/04/2020. Disponível em: http://www.stf.jus.br/arquivo/cms/noticiaNoticiaStf/anexo/ADI6387MC.pdf. Acesso em: 25 nov. 2020.

BRASIL. Supremo Tribunal Federal. *ADI nº 6.387/DF*. Rel. Min. Rosa Weber. Data de Publicação: 12/11/2020. Disponível em: http://portal.stf.jus.br/processos/detalhe.asp?incidente=5895165. Acesso em: 12 abr. 2021.

BRASIL. Supremo Tribunal Federal. *ADI nº 6.389/DF*. Relator: Min. Rosa Weber. Data de Julgamento: 26/11/2020. Data de Publicação: 30/11/2020. Disponível em: http://portal.stf.jus.br/processos/detalhe.asp?incidente=5895168. Acesso em: 24 fev. 2021.

BRASIL. Tribunal Regional do Trabalho da 1ª Região. *Mandado de Segurança Cível — Processo nº 0103519-41.2020.5.01.0000.* Julgado em: 29/04/2021.

BRASIL. Tribunal Regional do Trabalho da 9º Região. *Rito Sumaríssimo nº 0000335-45.2020.5.09.0130*. Relator: Leonardo Vieira Wandelli. Concluso em 06/11/2020.

BRAUN, I. *High-Risk Citizens*. 2018. Disponível em: https://algorithmwatch.org/en/story/high-risk-citizens/. Acesso em: 28 jul. 2020.

BRKAN, M.; BONNET, G. Legal and Technical Feasibility of the GDPR's Quest for Explanation of Algorithmic Decisions: of Black Boxes, White Boxes and Fata Morganas. *European Journal of Risk Regulation*, v. 11, n. 1, 2020.

BUCHANAN, B. G. A. (Very) Brief History of Artificial Intelligence. *AI Magazine*, [*S. l.*], v. 26, n. 4, 2005. Disponível em: https://doi.org/10.1609/aimag.v26i4.1848. Acesso em: 21 jun. 2021.

BURNS, K.; BECHARA, A. Decision making and free will: a neuroscience perspective. *Behavioral Sciences & the Law*, [*S. l.*], v. 25, n. 2, p. 263–280, 2007. Disponível em: https://doi.org/10.1002/bsl.751. Acesso em: 28 jun. 2021.

BURRELL, J. *How the Machine "Thinks:"* Understanding Opacity in Machine Learning Algorithms. Rochester, NY: Social Science Research Network, 2015. Disponível em: https://papers.ssrn.com/abstract=2660674. Acesso em: 27 maio. 2020.

CASEY, B.; FARHANGI, A.; VOGL, R. Rethinking Explainable Machines: The GDPR's "Right to Explanation" Debate and the Rise of Algorithmic Audits in Enterprise. In: *Berkeley Technology Law Journal*, v. 34, p. 145-189, 2018. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3143325. Acesso em: 19 maio 2019.

CAVOUKIAN, A. Privacy by Design: Origins, meaning, and prospects for assuring privacy and trust in the information era. *In*: YEE, G. O. M. *Privacy Protection Measures and Technologies in Business Organizations*: Aspects and Standards, [*S. l.*], IGI Global, 2012. p. 170–208. Disponível em: https://doi.org/10.4018/978-1-61350-501-4.ch007. Acesso em: 21 jun. 2021.

CENTRE FOR INFORMATION POLICY LEADERSHIP. O papel da Autoridade Nacional de Proteção de Dados (ANPD) conforme a nova Lei Geral de Proteção de Dados Pessoais (LGPD). *CIPL*, 17 abr. 2020.

CHAO, G. Human-Computer Interaction: Process and Principles of Human-Computer Interface *Design*. *In*: *2009 International Conference on Computer and Automation Engineering*, [*S. l.*]: 2009. p. 230–233. Disponível em: https://doi.org/10.1109/ICCAE.2009.23. Acesso em: 21 jun. 2021.

CHEN, W.; BARRY, W. Charting Digital Divides: Comparing Socioeconomic, Gender, Life Stage, and Rural-Urban Internet Access and Use in Five Countries. *In*: DUTTON, W. H.; KAHIN, B.; O'CALLAGHAN, R.; WYCKOFF, A. W. *Transforming Enterprise:* The Economic and Social Implications of Information Technology. MITP, 2004. p. 467–497. E-book.

CHENEY-LIPPOLD, J. *We are data*: algorithms and the making of our digital selves. New York: New York University Press, 2017.

CHOPRA, S.; WHITE, L. F. *A Legal Theory for Autonomous Artificial Agents*. [*S. l.*]: University of Michigan Press, 2011. Disponível em: https://www.jstor.org/stable/10.3998/mpub.356801. Acesso em: 27 maio. 2020.

CINTRA, A. C. de A.; GRINOVER, A. P.; DINAMARCO, C. R. *Teoria Geral do Processo*. 27. ed. São Paulo: Malheiros, 2011.

CITRON, D. K. Technological Due Process*. Washington University Law Review*, v. 85, p. 1249-1313, 2007. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1012360. Acesso em: 21 jun. 2021.

CITRON, D. K.; PASQUALE, F. The Scored Society: Due Process for Automated Predictions. *Washington Law Review*, v. 89, 2014. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2376209. Acesso em: 2 abr. 2021.

CNIL. *Privacy Impact Assessment Methodology.* [*S. l.*]: CNIL, 2018. Disponível em: https://www.cnil.fr/sites/default/files/atoms/files/cnil-pia-1-en-methodology.pdf. Acesso em: 9 jun. 2020.

COALIZÃO DIREITOS NA REDE. Coalizão Direitos na Rede repudia os 9 vetos de Bolsonaro à lei que cria a Autoridade Nacional de Proteção de Dados. *Medium*, 9 jul. 2019. Disponível em: https://cdr-br.medium.com/coaliz%C3%A3o-direitos-na-rede-repudia-os-9-vetos-de-bolsonaro-%C3%A0-lei-que-cria-a-autoridade-nacional-de-ee536f6baeb. Acesso em: 28 abr. 2021:

COELHO, F. U. *Curso de Direito Comercial.* Direito de Empresa. São Paulo: Revista dos Tribunais, 2018. 2 v.

COMPETITION & MARKETS AUTHORITY. *Online platforms and digital advertising*: market study interim report. [*S. l.*]: CMA, 2019. Disponível em: https://assets.publishing.service.gov.uk/media/5dfa0580ed915d0933009761/Interim_report.pdf. Acesso em: 26 abr. 2021.

CORBETT-DAVIES, S. *et al*. Algorithmic decision making and the cost of fairness. *arXiv*, 2017. Disponível em: http://arxiv.org/abs/1701.08230. Acesso em: 14 set. 2020.

CRAWFORD, K.; SCHULTZ, J. *Big Data and Due Process:* Toward a Framework to Redress Predictive Privacy Harms. Rochester, NY: Social Science Research Network, 2013. Disponível em: https://papers.ssrn.com/abstract=2325784. Acesso em: 27 maio 2020.

CUEVA, R. V. B. A proteção dos dados pessoais na jurisprudência do Superior Tribunal de Justiça. *In:* TEPEDINO, G.; FRAZÃO, A.; OLIVA, M. D. *Lei Geral de Proteção de Dados Pessoais e suas repercussões no Direito Brasileiro*. 2. ed. São Paulo: Thomson Reuters, 2020.

DATA PRIVACY BRASIL. *Memória da LGPD*: como a lei mudou desde 2010? *Observatório da Privacidade*, 2019. Disponível em: https://www.observatorioprivacidade.com.br/memoria/como-a-lei-mudou-desde-2010/. Acesso em: 11 abr. 2021.

DAVIES, R. S. Understanding Technology Literacy: A Framework for Evaluating Educational Technology Integration. *TechTrends*, [*S. l.*], v. 55, n. 5, 2011. Disponível em: https://doi.org/10.1007/s11528-011-0527-3. Acesso em: 21 jun. 2021.

DECIDIR. *In*: MICHAELIS. São Paulo: Editora Melhoramentos, 2021. Disponível em: https://michaelis.uol.com.br/busca?r=0&f=0&t=0&palavra=decidir. Acesso em: 10 jun. 2021.

DEMETZOU, K. Data Protection Impact Assessment: A tool for accountability and the

unclarified concept of ‘high risk’ in the General Data Protection Regulation. *Computer Law & Security Review*, [*S. l.*], v. 35, n. 6, p. 105342, 2019. Disponível em: https://doi.org/10.1016/j.clsr.2019.105342. Acesso em: 21 jun. 2021.

DEVORE, J. L. *Probabilidade e estatística*: para engenharia e ciências. Tradução: Joaquim Pinheiro Nunes da Silva. São Paulo: Cengage Learning, 2006.

DIAKOPOULOS, N. *Algorithmic accountability reporting:* on the investigation of black boxes. [*S. l.*]: Tow Center for Digital Journalism, 2014. Disponível em: https://academiccommons.columbia.edu/doi/10.7916/D8ZK5TW2. Acesso em: 21 jun. 2021.

DIGITAL rights alliance file legal complaints across Europe against facial recognition Company Clearview AI. *Noyb*, 26 maio 2021. Disponível em: https://noyb.eu/en/digital-rights-alliance-file-legal-complaints-against-facial-recognition-company-clearview-ai. Acesso em: 29 maio 2021.

DONEDA, D. A Autoridade Nacional de Proteção de Dados e o Conselho Nacional de Proteção de Dados. *In*: MENDES, L. S.; DONEDA, D.; SARLET, I. W.; RODRIGUES JÚNIOR, O. L.; BIONI, B. (Coords.). *Tratado de Proteção de Dados Pessoais*. Rio de Janeiro: Forense, 2021.

DONEDA, D. *Da privacidade à proteção de dados pessoais*. 7. ed. São Paulo: Revista dos Tribunais, 2019.

DONEDA, D. *Da privacidade à proteção de dados pessoais:* fundamentos da Lei Geral de Proteção de dados. 2. ed. São Paulo: Thomson Reuters Brasil, 2019.

DONEDA, D.; ALMEIDA, V. What Is Algorithm Governance? *IEEE Internet Computing*, [*S. l.*], v. 20, p. 60–63, 2016. Disponível em: https://doi.org/10.1109/MIC.2016.79. Acesso em: 21 jun. 2021.

DONEDA, Danilo. Current Judicial and Administrative Issues of Consumer Data Protection in Brazil. *In*: METZ, R.; BINDING, J.; HAIFEND, P. (Eds.). *Consumer Protection in Brazil, China and Germany:* a comparative study. Göttingen: Göttingen University Press, 2016. Disponível em: https://univerlag.uni-goettingen.de/handle/3/isbn-978-3-86395-236-5. Acesso em: 08 abr. 2021.

DOSHI-VELEZ, F.; KIM, B. Towards A Rigorous Science of Interpretable Machine Learning. *arXiv*, 2017. Disponível em: http://arxiv.org/abs/1702.08608. Acesso em: 30 jun. 2020.

DUARTE, M. de F.; BRAGA, C. P. *Propriedade Intelectual*. 1.· ed. [*S. l.*]: Sagah, 2018.

EDPS. *Survey on Data Protection Impact Assessments under Article 39 of the Regulation*. [*S. l.*]: EDPS, 2020. Disponível em: https://edps.europa.eu/sites/edp/files/publication/20-07-06_edps_dpias_survey_en.pdf. Acesso em: 9 jun. 2020.

EDWARDS, L.; VEALE, M. Slave to the Algorithm: Why a Right to an Explanation Is

Probably Not the Remedy You Are Looking for. *Duke Law & Technology Review*, v. 16, p. 18-84, 2017. Disponível em: https://osf.io/preprints/lawarxiv/97upg/. Acesso em: 7 nov. 2020.

EHSAN, U.; RIEDL, M. O. *On Design and Evaluation of Human-centered Explainable AI systems*. Glasgow, 2019. Disponível em: https://www.cc.gatech.edu/~riedl/pubs/ehsan-chi-hcml19.pdf. Acesso em: 22 set. 2020.

EKKER. *Uber drivers demand access to their personal data*. Disponível em: https://ekker.legal/2020/07/19/uber-drivers-demand-access-to-their-personal-data/. Acesso em: 1 jun. 2021.

ESTADOS UNIDOS DA AMÉRICA. Constituição dos Estados Unidos da América. 1787.

ESTADOS UNIDOS DA AMÉRICA. *Privacy Act, 88 Stat. 1896 Public Law*. Washington D.C., 31 dez. 1974.

ESTADOS UNIDOS DA AMÉRICA. *Electronic Communications Privacy Act, 18 U.S.C. §2510 e ss., Public Law*. Washington D.C., 21 out. 1986.

ESTADOS UNIDOS DA AMÉRICA. *Health Insurance Portability and Accountability Act*. *110 Stat. 1936, Public Law*. Washington D.C., 21 ago. 1996.

ESTADOS UNIDOS DA AMÉRICA. *Children's Online Privacy Protection Act, 15 U.S.C*. *§6501-6506., Public Law*. Washington D.C., 21 out. 1998.

ESTADOS UNIDOS DA AMÉRICA. *H.R.4368 — 116th Congress (2019-2020)*: Justice in Forensic Algorithms Act of 2019. Washington D.C., 10 fev. 2019. Disponível em: https://www.congress.gov/bill/116th-congress/house-bill/4368/text. Acesso em: 24 jul. 2020.

ESTADOS UNIDOS DA AMÉRICA. *H.R. 2231 — 116th Congress (2019-2020):* Algorithmic Accountability Act of 2019. 4 nov. 2019. Disponível em: https://www.congress.gov/bill/116th-congress/house-bill/2231/all-info. Acesso em: 24 jul. 2020.

EUBANKS, Virginia. *Automating inequality:* how high-tech tools profile, police, and punish the poor. New York: St Martin's Press, 2018.

EUROPEAN DATA PROTECTION BOARD. *ARTICLE 29 Newsroom* — Guidelines on the right to "data portability". Disponível em: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611233. Acesso em: 26 abr. 2021.

EUROPEAN DATA PROTECTION BOARD. *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 679/2016*. Bruxelas: European Commission, 2016. Disponível em: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053. Acesso

em: 21 jun. 2021.

EUROPEAN DATA PROTECTION BOARD. *Guidelines on Data Protection Impact Assessment (DPIA) (wp248rev.01)*. [*S. l.*]: EDPB, 2017. Disponível em: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611236. Acesso em: 17 nov. 2020.

EUROPEAN DATA PROTECTION BOARD. *Guidelines 2/2019 on the processing of personal data under Article 6(1)(b) GDPR in the context of the provision of online services to data subjects*. [S. l.]: EDPB, 2019. Disponível em: https://privacyblogfullservice.huntonwilliamsblogs.com/wp-content/uploads/sites/28/2019/04/edpb_draft_guidelines-art_6-1-b-final_public_consultation_version_en.pdf. Acesso em: 26 abr. 2021.

EUROPEAN DATA PROTECTION BOARD. *Guidelines 07/2020 on the concepts of controller and processor in the GDPR*. [*S. l.*]: EDPB, 2020. Disponível em: https://edpb.europa.eu/sites/default/files/consultation/edpb_guidelines_202007_controllerprocessor_en.pdf. Acesso em: 3 maio 2021.

EUROPEAN UNION AGENCY FOR FUNDAMENTAL RIGHTS. *Data Protection in the European Union*: the role of National Data Protection Authorities. Bruxelas: FRA, 2010. Disponível em: https://fra.europa.eu/sites/default/files/fra_uploads/815-Data-protection_en.pdf. Acesso em: 16 maio 2021.

FAMÍLIAS sem acesso à internet não conseguem usar o dinheiro do auxílio emergencial. *Jornal Nacional*, 9 abr. 2021. Disponível em: https://g1.globo.com/jornal-nacional/noticia/2021/04/09/familias-sem-acesso-a-internet-nao-conseguem-usar-o-dinheiro-do-auxilio-emergencial.ghtml. Acesso em: 22 maio 2021.

FEDERAL TRADE COMMISSION. A Brief Overview of the Federal Trade Commission's Investigative, Law Enforcement, and Rulemaking Authority. *FTO*, out. 2019. Disponível em: https://www.ftc.gov/about-ftc/what-we-do/enforcement-authority. Acesso em: 22 mar. 2021.

FEDERAL TRADE COMMISSION. *Big Data:* A Tool for Inclusion or Exclusion? Understanding the Issues. [*S. l.*]: FTC, 2016. Disponível em: https://www.ftc.gov/reports/big-data-tool-inclusion-or-exclusion-understanding-issues-ftc-report. Acesso em 23 jul. 2020.

FEDERAL TRADE COMMISSION. FTC Policy Statement on Deception. *FTC*, Washington, D.C., 14 out. 1983. Disponível em: https://www.ftc.gov/system/files/documents/public_statements/410531/831014deceptionstmt.pdf. Acesso em: 22 mar. 2021.

FELTON, R. Michigan unemployment agency made 20,000 false fraud accusations – report. *The Guardian*, Detroit, 18 dez. 2016. Disponível em: http://www.theguardian.com/us-news/2016/dec/18/michigan-unemployment-agency-fraud-accusations. Acesso em: 26 maio 2021.

FERRAZ JÚNIOR, T. S. *Introdução ao estudo do direito técnica, decisão, dominação*.

São Paulo: Atlas, 2013.

FISCHER, G. User Modeling in Human–Computer Interaction. *User Modeling and User-Adapted Interaction*, v. 11, n. 1, p. 65–86, 2001. Disponível em: https://doi.org/10.1023/A:1011145532042. Acesso em: 21 jun. 2021.

FLORÊNCIO, J. A. *Proteção de dados na cultura do algoritmo*. 2019. 320 f. Tese (Doutorado em Direito) — Programa de Estudos Pós-Graduados em Direito, Pontifícia Universidade Católica de São Paulo, São Paulo, 2019.

FLORIDI, L. The Informational Nature of Personal Identity*. Minds and Machines*, v. 21, n. 4, p. 549–566, 2011.

FLORIDI, L. Open Data, Data Protection, and Group Privacy. *Philosophy & Technologys.* v. 27, n. 1, p. 1–3, 2014. Disponível em: https://doi.org/10.1007/s13347-014-0157-8. Acesso em: 21 jun. 2021.

FORSYTH, D. R. Conflict. *In:* FORSYTH, D. R. *Group dynamics*. 5. ed. Belmont: CA: Wadsworth, Cengage Learning, 2006.

FRAZÃO, A. Nova LGPD: direitos dos titulares de dados pessoais. *Jota*, 24 out. 2018. Disponível em: https://www.jota.info/opiniao-e-analise/colunas/constituicao-empresa-e-mercado/nova-lgpd-direitos-dos-titulares-de-dados-pessoais-24102018. Acesso em: 06 jan. 2021.

FRAZÃO, A. O direito à explicação e à oposição diante de decisões totalmente automatizadas. *Jota*, 05 dez. 2018. Disponível em: https://www.jota.info/opiniao-e-analise/colunas/constituicao-empresa-e-mercado/o-direito-a-explicacao-e-a-oposicao-diante-de-decisoes-totalmente-automatizadas-05122018. Acesso em: 11 abr. 2021.

FRAZÃO, A. Fundamentos da proteção dos dados pessoais: noções introdutórias para a compreensão da importância da Lei Geral de Proteção de Dados. *In*: TEPEDINO, Gustavo; FRAZÃO, Ana; OLIVA, Milena Donato. *Lei Geral de Proteção de Dados Pessoais e suas repercussões no Direito Brasileiro*. 2. ed. São Paulo: Thomson Reuters, 2020. p. 23-52.

FRIEDLER, S. A.; SCHEIDEGGER, C.; VENKATASUBRAMANIAN, S. On the (im)possibility of fairness. *arXiv*, 2016. Disponível em: http://arxiv.org/abs/1609.07236. Acesso em: 18 jun. 2020.

FROOMKIN, A. M. Regulating Mass Surveillance as Privacy Pollution: Learning from Environmental Impact Statements. *U. Ill. L. Rev. 1713*, [*S. l.*], v. 2015, 2015.

FRY, H. *Hello World*: How to be Human in the Age of the Machine. London New York Toronto Sidney Auckland: Doubleday, 2018.

GDPRHUB. Welcome to GDPRhub. Disponível em: https://gdprhub.eu/index.php?title=Welcome_to_GDPRhub. Acesso em: 21 jun. 2021.

GELLERT, R. Understanding data protection as risk regulation. *Journal of Internet Law*, [*S. l.*], p. 3–15, 2015.

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. Disponível em: https://doi.org/10.1016/j.clsr.2017.12.003. Acesso em: 21 jun. 2021.

GELLMAN, R. *Fair Information Practices*: A Basic History. v. 2.19. Rochester, NY: Social Science Research Network, 2019. Disponível em: https://papers.ssrn.com/abstract=2415020. Acesso em: 11 jun. 2020.

GIDDENS, A. *Modernidade e identidade*. Rio de Janeiro: Zahar, 2002.

GILLESPIE, T. Algorithm*. In:* PETERS, B. *Digital Keywords*. Princeton: Princeton University Press, 2016. p. 18-30.

GOMES, M. C. Relatório de impacto à proteção de dados. Uma breve análise da sua definição e papel na LGPD. *Revista do Advogado*, [*S. l.*], v. 39, n. 144, p. 174–183, 2019.

GOODMAN, B.; FLAXMAN, S. European Union regulations on algorithmic decision-making and a “right to explanation”. *AI Magazine*, [*S. l.*], v. 38, n. 3, p. 50–57, 2017. Disponível em: https://arxiv.org/abs/1606.08813. Acesso em: 15 dez. 2020.

GOOGLE INC. *ML Practicum*: Image Classification. Google Developers. Disponível em: https://developers.google.com/machine-learning/practica/image-classification. Acesso em: 22 maio 2021.

GREENLEAF, G. The Influence of European Data Privacy Standards Outside Europe: Implications for Globalisation of Convention 108. *International Data Privacy Law*, [*S. l.*], v. 2, 2011.

GUAGNIN, D. *et al.* Introduction. *In*: GUAGNIN, D. *et al.* (Org.). *Managing Privacy through Accountability*. London: Palgrave Macmillan UK, 2012. p. 1–14. E-book.

GUIDI, G. B. de C. *Modelos regulatórios para proteção de dados pessoais*. Rio de Janeiro: Instituto de Tecnologia e Sociedade do Rio, 2017. Disponível em: https://itsrio.org/wp-content/uploads/2017/03/Guilherme-Guidi-V-revisado.pdf. Acesso em: 22 mar. 2021.

GUIDOTTI, R. *et al.* Local Rule-Based Explanations of Black Box Decision Systems. *arXiv*, 2018. Disponível em: http://arxiv.org/abs/1805.10820. Acesso em: 30 jun. 2020.

GUNST, H. *The Right to Explanation and the Right to Secrecy* – Reconciling Data Protection and Trade Secret Rights in Automated Decision-making. 2017. Dissertation (Masters Thesis in Law) – Faculty of Law, University of Helsinki, Helsinki, 2017. Disponível em: http://hdl.handle.net/10138/231948. Acesso em: 20 jun. 2021.

HANSEN, H. K.; FKYVERBOM, M. The politics of transparency and the calibration of knowledge in the digital age. *Organization*, v. 22, n. 6, p. 872–889, 2015.

HANSSON, S. O. Decision Theory: An Overview. *In*: LOVRIC, M. (Org.). *International Encyclopedia of Statistical Science*. Berlin, Heidelberg: Springer, 2011. p. 349–355. Disponível em: https://doi.org/10.1007/978-3-642-04898-2_22. Acesso em: 14 jun. 2021.

HILL, K. How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did. *Forbes*, 16 fev. 2012. Disponível em: https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/. Acesso em: 27 ago. 2020.

HOME Office drops “Racist” algorithm from Visa Decisions. *BBC News*, [*S. l.*], 4 ago. 2020. Disponível em: https://www.bbc.com/news/technology-53650758. Acesso em: 27 ago. 2020.

HOOD, C.; HEAD, D. *Transparency*: The Key to Better Governance? Oxford: Oxford University Press, British Academy, 2006.

HOOFNAGLE, C. J. *The Origin of Fair Information Practices:* Archive of the Meetings of the Secretary’s Advisory Committee on Automated Personal Data Systems (SACAPDS). Berkeley Law, 16 jul. 2014. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2466418. Acesso em: 21 jun. 2021.

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*, v. 1, n. 2, dez. 2020, p. 90. Disponível em: https://revista.internetlab.org.br/wp-content/uploads/2020/12/Tomada-de-Decisa%CC%83o-Automatizada.pdf. Acesso em: 11 abr. 2021.

INFORMATION COMMISSIONER'S OFFICE. *Big data, artificial intelligence, machine learning and data protection*. 2017. Disponível em: https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf. Acesso em: 28 jul. 2020.

INFORMATION COMMISSIONER’S OFFICE; THE ALAN TURING INSTITUTE. ICO and Turing consultation on Explaining AI decisions guidance. *ICO*, 24 jan. 2020. Disponível em: https://ico.org.uk/about-the-ico/ico-and-stakeholder-consultations/ico-and-the-turing-consultation-on-explaining-ai-decisions-guidance/. Acesso em: 16 abr. 2021.

INFORMATION COMMISSIONER’S OFFICE. *Guidance on the AI auditing framework*. [*S. l.*]: ICO, 2020. Disponível em: https://ico.org.uk/media/about-the-ico/consultations/2617219/guidance-on-the-ai-auditing-framework-draft-for-consultation.pdf. Acesso em: 1 ago. 2020.

INSTITUTO DE REFERÊNCIA EM INTERNET E SOCIEDADE. Right to Explanation and Artificial Intelligence. *IRIS BH*, 17 jun. 2019. Disponível em: https://irisbh.com.br/en/right-to-explanation-and-artificial intelligence-brief-considerations-on-the-european-debate/. Acesso em: 7 nov. 2020.

INSTITUTO IGARAPÉ. *Reconhecimento Facial no Brasil*. INFOGRÁFICO RECONHECIMENTO FACIAL NO BRASIL — INSTITUTO IGARAPÉ. Instituto Igarapé, 2020. Disponível em: https://igarape.org.br/infografico-reconhecimento-facial-no-brasil/. Acesso em: 29 maio 2021.

INTERNETLAB. O que está em jogo no debate sobre dados pessoais no Brasil? Relatório final sobre o debate público promovido pelo Ministério da Justiça sobre o Anteprojeto de Lei de Proteção de Dados Pessoais. São Paulo: InternetLab, 2016. Disponível em: https://www.internetlab.org.br/wp-content/uploads/2016/05/reporta_apl_dados_pessoais_final.pdf. Acesso em: 4 maio 2021.

INTRODUCING A NEW FEATURE: DRIVING HOURS LIMIT. *UBER*, 2018. Disponível em: https://www.uber.com/en-ZA/blog/driving-hours-limit/. Acesso em: 5 jun. 2021.

ISA. WHAT IS AUTOMATION? Disponível em: https://www.isa.org/about-isa/what-is-automation. Acesso em: 14 jun. 2021.

JAMES, W. What pragmatism means. *In:* MENAND, L. *Pragmatism: A Reader*. New York: Random House, 1997. p. 93–111.

JANSSEN, J. H. N. *The right to explanation:* means for 'white-boxing' the black-box?: research into the ability of the 'right to explanation' about decisions based solely on automated decision-making of Articles 13(2)(f), 14(2)(g), 15(1)(h) and 22(3) of the General Data Protection Regulation, as well as of current explanation methods, to solve the legal problems arising from algorithmic decision-making. Dissertação (Masters Thesis in Law and Technology) — Universiteit van Tilburg, Tilburg, 2019. Disponível em: https://tilburguniversity.on.worldcat.org/search?queryString=scr.uvt.nl:8107234. Acesso em: 20 jun. 2021.

JILLSON, E. Aiming for truth, fairness, and equity in your company's use of AI. *Federal Trade Commision*, 19 abr. 2021. Disponível em: https://www.ftc.gov/news-events/blogs/business-blog/2021/04/aiming-truth-fairness-equity-your-companys-use-ai. Acesso em: 22 jun. 2021.

KAIYAN, N. Exploratory study of implicit theories in human computer interaction. *Proceedings Sixth Australian Conference on Computer-Human Interaction*. [*S. l.: s. n.*], 1996. p. 338–339. Disponível em: https://doi.org/10.1109/OZCHI.1996.560158. Acesso em: 21 jun. 2021.

KAK, A. Regulating Biometrics: Global Approaches and Urgent Questions. *AI Now Institute*, set. 2020. Disponível em: https://ainowinstitute.org/regulatingbiometrics.pdf. Acesso em: 20 abr. 2021.

KAMINSKI, M. E. The Right to Explanation, Explained. *Berkeley Technology Law Journal*, v. 34, n. 189, 2019. Disponível em: https://papers.ssrn.com/abstract=3196985. Acesso em: 27 maio. 2020.

KAMINSKI, M. E.; MALGIERI, G. Algorithmic impact assessments under the GDPR: producing multi-layered explanations. *International Data Privacy Law*, [*S. l.*], n. ipaa020, 2020. Disponível em: https://doi.org/10.1093/idpl/ipaa020. Acesso em: 29 mar. 2021.

KATYAL, S. *Private Accountability in the Age of Artificial Intelligence*. Rochester, NY: Social Science Research Network, 2019. SSRN Scholarly Paper. Disponível em: https://papers.ssrn.com/abstract=3309397. Acesso em: 21 abr. 2021.

KELLER, C. I. *Regulação nacional de serviços na Internet*: exceção, legitimidade e o papel do Estado. Rio de Janeiro: Lumen Juris, 2019.

KELSEN, H. *Pure theory of law*. Union, N.J: Lawbook Exchange, 1967.

KLEIN, A. Reducing bias in AI-based financial services. *BROOKINGS*. 10 jul. 2020. Disponível em: https://www.brookings.edu/research/reducing-bias-in-ai-based-financial-services/. Acesso em: 1 set. 2020.

KOLKMAN, D. The (in)credibility of algorithmic models to non-experts. *Information, Communication & Society*, v. 0, n. 0, p. 1–17, 2020. Disponível em: https://doi.org/10.1080/1369118X.2020.1761860. Acesso em: 28 jun. 2021.

LARSON, J.; MATTU, S.; KIRCHNER, L.; ANGWIN, J. How We Analyzed the COMPAS Recidivism Algorithm. *ProPublica*, 23 maio 2016. Disponível em: https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm. Acesso em: 31 maio 2021.

LESLIE, D. Understanding bias in facial recognition technologies. *arXiv*, 2020. Disponível em: https://doi.org/10.5281/zenodo.4050457. Acesso em: 29 maio 2021.

LIAO, Q. V.; GRUEN, D.; MILLER, S. Questioning the AI: Informing *Design* Practices for Explainable AI User Experiences. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems,* [*S. l.*], p. 1–15, abr. 2020. Disponível em: https://doi.org/10.1145/3313831.3376590. Acesso em; 21 jun. 2021.

LODGE, J. Transparency and Democratic Legitimacy. *JCMS: Journal of Commom Market Studie***s**, [*S. l.*], 1994. Disponível em: https://doi.org/10.1111/j.1468-5965.1994.tb00501.x. Acesso em: 4 dez. 2020.

LOSH, S. C. Generation versus aging, and education, occupation, gender and ethnicity effects in U.S. digital divides. *In: 2009 Atlanta Conference on Science and Innovation Policy,* Atlanta, 2009. p. 1–8. Disponível em: https://doi.org/10.1109/ACSIP.2009.5367820. Acesso em: 21 jun. 2021.

LYNSKEY, O. *The Foundations of EU Data Protection Law*. Oxford: Oxford University Press, 2015.

MACCARTHY, M. Standards of Fairness for Disparate Impact Assessment of Big Data Algorithms. *Cumberland Law Review*, [*S. l.*], v. 48, n. 1, p. 67–148, 2017.

MACHADO, C. C. V. Cidade dos algoritmos: A Ética da Informação nas Cidades Inteligentes. *Instituto de Tecnologia e Sociedade do Rio — ITS Rio*, mar. 2018. Disponível em: https://itsrio.org/wp-content/uploads/2018/03/caio_machado_etica.pdf. Acesso em: 19 abr. 2021.

MAHIEU, R.; AUSLOOS, J. *Recognising and Enabling the Collective Dimension of the GDPR and the Right of Access*. *LawArXiv*, 2 jul. 2020. Disponível em: https://doi.org/10.31228/osf.io/b5dwm. Acesso em: 9 dez. 2020.

MAJOR EU-US data protection agreement struck down. *BBC News*, [*S. l.*], 2020. Disponível em: https://www.bbc.com/news/technology-53418898 . Acesso em: 24 jul. 2020.

MANTELERO, A. AI and Big Data: A blueprint for a human rights, social and ethical impact assessment. *Computer Law & Security Review,* [*S. l.*], v. 34, n. 4, p. 754–772, 2018. Disponível em: https://doi.org/10.1016/j.clsr.2018.05.017. Acesso em: 21 jun. 2021.

MARTINEZ, A. Considering AI In Hiring? As Its Use Grows, So Do The Legal Implications For Employers. *Forbes*, 5 dez. 2020. Disponível em: https://www.forbes.com/sites/alonzomartinez/2019/12/05/considering-ai-in-hiring-as-its-use-grows-so-do-the-legal-implications-for-employers/#45c4dca77d47. Acesso em: 1 set. 2020.

MASSÉ, E. *Two Years Under the EU GDPR:* An Implementation Progress Report. [*S. l.*]: AccessNow, 2020. Disponível em: https://www.accessnow.org/cms/assets/uploads/2020/05/Two-Years-Under-GDPR.pdf. Acesso em: 16 maio 2021.

MATTIUZO, M. Discriminação algorítmica: reflexões no contexto da Lei Geral de Proteção de Dados Pessoais. *In*: DONEDA, D.; MENDES, L. S.; CUEVA, R. V. B. (Coords.). *Lei Geral de Proteção de Dados (Lei nº 13.709/2018) — A caminho da efetividade:* contribuições para a implementação da LGPD. São Paulo: Thomsom Reuters, 2020.

MAYER-SCHOENBERGER, V. Generational development of data protection in Europe. *In*: AGRE, P. E.; ROTENBERG, M. (Eds.). *Technology and Privacy:* The New Landscape. Cambridge: MIT Press, 1998. p. 219-241.

MAYER-SCHÖNBERGER, V. *Delete*. [*S. l.*]: Princeton University Press, 2009. Disponível em: www.jstor.org/stable/j.ctt7t09g . Acesso em: 27 jul. 2020.

MENDES, L. S. *Privacidade, proteção de dados e defesa do consumidor*: linhas gerais de um novo direito fundamental. São Paulo: Saraiva, 2014.

MENDES, L. S.; RODRIGUES JÚNIOR, O. L.; FONSECA, G. C. S. da. O Supremo Tribunal Federal e a Proteção Constitucional dos Dados Pessoais: Rumo a um Direito Fundamental Autônomo. *In*: MENDES, L. S.; DONEDA, D.; SARLET, I. W.; RODRIGUES JÚNIOR, O. L.; BIONI, B. (Coords.). *Tratado de Proteção de Dados Pessoais*. Rio de Janeiro: Forense, 2021.

MICHENER, G.; BERSCH, K. Identifying Transparency. *Inf. Polity*, v. 18, n. 3, p. 233-242, 2013. Disponível em: https://doi.org/10.3233/IP-130299. Acesso em: 21 jun.2021.

MILLER, J. M. *Dignity as a New Framework, Replacing the Right to Privacy*. Rochester, NY: Social Science Research Network, 2008. Disponível em: https://papers.ssrn.com/abstract=1127986. Acesso em: 27 maio. 2020.

MILLER, T. Explanation in Artificial Intelligence: Insights from the Social Sciences. *ArXiv,* [*S. l.*], 2018. Disponível em: http://arxiv.org/abs/1706.07269. Acesso em: 14 ago. 2020.

MIT MEDIA LAB. *The Moral Machine*. Disponível em: http://moralmachine.mit.edu. Acesso em: 29 abr. 2020.

MITTELSTADT, B. D. *et al*. The ethics of algorithms: Mapping the debate. *Big Data & Society*, [*S. l.*], v. 3, n. 2, 2016. Disponível em: https://doi.org/10.1177/2053951716679679. Acesso em: 20 maio 2020.

MITTELSTADT, B.; RUSSELL, C.; WACHTER, S. Explaining Explanations in AI. *In*: THE CONFERENCE, 2019, Atlanta, GA, USA. *Proceedings of the Conference on Fairness, Accountability, and Transparency — FAT'19*. Atlanta, GA, USA: ACM Press, 2019. p. 279–288. Disponível em: https://doi.org/10.1145/3287560.3287574. Acesso em: 27 maio 2020.

MOLNAR, C. *Interpretable Machine Learning*. [*S. l.*]: [*s. n.*], 2020. *E-book*.

MONTEIRO, R. L. Lei Geral de Proteção de Dados do Brasil: uma análise detalhada. *Jota*, 14 jul. 2018. Disponível em: https://www.jota.info/opiniao-e-analise/colunas/agenda-da-privacidade-e-da-protecao-de-dados/lgpd-analise-detalhada-14072018. Acesso em: 07 abr. 2021.

MONTEIRO, R. L. *Existe um direito à explicação na Lei Geral de Proteção de Dados do Brasil?* Rio de Janeiro: Instituto Igarapé, dez. 2018. Artigo Estratégico nº 39. Disponível em: https://igarape.org.br/wp-content/uploads/2018/12/Existe-um-direito-a-explicacao-na-Lei-Geral-de-Protecao-de-Dados-no-Brasil.pdf. Acesso em: 21 jun. 2021.

MONTEIRO, R. L.; CRUZ, S. N. e. Direitos dos titulares na LGPD: fundamentos, limites e aspectos práticos. *In*: FRANCOSKI, D. de S. L.; TASSO, F. A. (Coords.). *A Lei Geral de Proteção de Dados Pessoais*: Aspectos Práticos e Teóricos Relevantes no Setor Público e Privado. São Paulo: Thomsom Reuters, abr. 2021.

MOOR, J. Are There Decisions Computers Should Never Make. *Nature and System*, [*S. l.*], v. 1, 1985.

MULGAN, R. Issues of Accountability. *In*: MULGAN, R. (Org.). *Holding Power to Account*: Accountability in Modern Democracies. London: Palgrave Macmillan UK, 2003. p. 1–35. Disponível em: https://doi.org/10.1057/9781403943835_1. Acesso em: 13 nov. 2020.

MULHOLLAND, C.; FRAJHOF, I. Z. Inteligência Artificial e a Lei Geral 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. (Coords.). *Inteligência Artificial e Direito:* Ética, Regulação e Responsabilidade. 2. ed. São Paulo: Revista dos Tribunais, 2020.

MULLIGAN, D. K.; KING, J. *Bridging the Gap between Privacy and Design*. Rochester, NY: Social Science Research Network, 2012. SSRN Scholarly Paper. Disponível em: https://papers.ssrn.com/abstract=2070401. Acesso em: 23 set. 2020.

MULLIGAN, D. K.; KLUTTZ, D.; KOHLI, N. *Shaping Our Tools*: Contestability as a Means to Promote Responsible Algorithmic Decision Making in the Professions. Rochester, NY: Social Science Research Network, 2019. SSRN Scholarly Paper. Disponível em: https://doi.org/10.2139/ssrn.3311894. Acesso em: 14 jan. 2021.

NEWMAN, A. P*. Protectors of Privacy*: Regulating Personal Data in the Global Economy. London: Cornell University Press, 2008.

NOJIRI, Sergio. Decisão judicial. In: CAMPILONGO, C. F.; GONZAGA, A. de A.; FREIRE, A. L. (Coords.). *Enciclopédia jurídica da PUC-SP*. Tomo: Teoria Geral e Filosofia do Direito. 1. ed. São Paulo: Pontifícia Universidade Católica de São Paulo, 2017. Disponível em: https://enciclopediajuridica.pucsp.br/verbete/57/edicao-1/decisao-judicial. Acesso em: 21 jun. 2021.

NUNES, Pablo. Levantamento revela que 90,5% dos presos por monitoramento facial no Brasil são negros. *The Intercept*, 21 nov. 2019. Disponível em: https://theintercept.com/2019/11/21/presos-monitoramento-facial-brasil-negros/. Acesso em: 31 maio 2021.

O'NEIL, C. *Weapons of Math Destruction:* How Big Data Increases Inequality and Threatens Democracy. New York, NY: Crown Books, 2016. *E-book*.

OAIC. *Privacy Impact Assessment Guide*. Sydney: OAIC, 2010. Disponível em: http://www.icb.org.au/out/?dlid=38156. Acesso em: 14 abr. 2021.

OII LONDON LECTURE. Show me your data and I'll tell you who you are. 30 out. 2018. Disponível em: https://www.oii.ox.ac.uk/videos/oii-london-lecture-show-me-your-data-and-ill-tell-you-who-you-are/. Acesso em: 21 jun. 2021.

PASQUALE, F. *The Black box society:* the secret algorithms that control money and information. First Harvard University Press paperback edition ed. Cambridge, Massachusetts; London, England: Harvard University Press, 2015.

PERKOWITZ, S. The Bias in the Machine: Facial Recognition Technology and Racial Disparities. *MIT Case Studies in Social and Ethical Responsibilities of Computing*, [*S. l.*], 2021. Disponível em: https://doi.org/10.21428/2c646de5.62272586. Acesso em: 29 maio 2021.

PETERSON, M. Decision Theory: An Introduction. *In*: LOVRIC, M. (Org.). *International*

Encyclopedia of Statistical Science. Berlin, Heidelberg: Springer, 2011. p. 346–349. Disponível em: https://doi.org/10.1007/978-3-642-04898-2_23. Acesso em: 14 jun. 2021.

PEW RESEARCH CENTER: INTERNET, SCIENCE & TECH. *Code-Dependent*: Pros and Cons of the Algorithm Age. [*S. l.*]: Pew Research Center, 2017.

PREFEITURA DE SÃO PAULO. Prefeitura de São Paulo anuncia parceria com Waze. *Cidade de São Paulo*, 20 set. 2017. Disponível em: http://www.capital.sp.gov.br/noticia/prefeitura-de-sao-paulo-anuncia-parceria-com-waze/. Acesso em: 19 abr. 2021.

PRIVACY INTERNATIONAL. *Data Is Power*: Profiling and Automated Decision-Making in GDPR. [s.d.]. Disponível em: http://privacyinternational.org/report/1718/data-power-profiling-and-automated-decision-making-gdpr. Acesso em: 31 jul. 2020.

PRIVACY INTERNATIONAL. *Data is Power: Profiling and Automated Decision-Making in GDPR*. PI, 9 abr. 2017. Disponível em: https://privacyinternational.org/sites/default/files/2018-04/Data%20Is%20Power-Profiling%20and%20Automated%20Decision-Making%20in%20GDPR.pdf. Acesso em: 7 nov. 2020.

QUELLE, C. *The Data Protection Impact Assessment, or*: How the General Data Protection Regulation May Still Come to Foster Ethically Responsible Data Processing. Rochester, NY: Social Science Research Network, 2015. SSRN Scholarly Paper. Disponível em: https://doi.org/10.2139/ssrn.2695398. Acesso em: 24 abr. 2021.

RAAB, C.; SZEKELY, I. Data Protection Authorities and Information Technology. *Computer Law and Security Review*, 1 jul. 2017. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2994898. Acesso em: 14 maio 2021.

RAJI, I. D. *et al.* Closing the AI accountability gap: defining an end-to-end framework for internal algorithmic auditing. *In*: FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency. New York, USA: Association for Computing Machinery, 2020. p. 33–44. Disponível em: https://doi.org/10.1145/3351095.3372873. Acesso em: 1 set. 2020.

RAMGE, T.; MAYER-SCHONBERGER, V. *Reinventing Capitalism in the Age of Big Data*. New York: Basic Books, 2018.

RAVICHANDRAN, D.; VASSILVITSKII, S. *Evaluation of Cohort Algorithms for the FLoC API*. [*S. l.*], [*s. n*.], 21 out. 2020. Disponível em: https://github.com/google/ads-privacy/blob/master/proposals/FLoC/FLOC-Whitepaper-Google.pdf.

RAWLS, J. *Uma Teoria da Justiça*. 3. ed. São Paulo: Martins Fontes, 2008.

RB. DEN HAAG. *ECLI:NL:RBDHA:2020:1013*, *Rechtbank Den Haag, C-09-585239-KG ZA 19-1221*. 2020. Disponível em:

https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBDHA:2020:1013. Acesso em: 28 jul. 2020.

RB. DEN HAAG. *ECLI:NL:RBDHA:2020:1878, Rechtbank Den Haag, C-09-550982-HA ZA 18-388*. 2020. Disponível em: https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBDHA:2020:1878. Acesso em: 28 jul. 2020.

REDISH, M.; MARSHALL, L. Adjudicatory Independence and the Values of Procedural Due Process. *Yale Law Journal*, [*S. l.*], v. 95, n. 3, 1986. Disponível em: https://digitalcommons.law.yale.edu/ylj/vol95/iss3/1. Acesso em: 21 jun. 2021.

REISMAN, D. *et al*. *Algorithm Impact Assessment:* a practical framework for public agency accountability. [*S. l.*]: AI Now Institute, 2018. Disponível em: https://ainowinstitute.org/aiareport2018.pdf. Acesso em: 21 abr. 2021.

RIBEIRO, M. T.; SINGH, S.; GUESTRIN, C. "Why Should I Trust You?": Explaining the Predictions of Any Classifier*. ArXiv*, 2016. Disponível em: http://arxiv.org/abs/1602.04938. Acesso em: 25 ago. 2020.

RIO GRANDE DO SUL. Tribunal de Justiça do Estado do Rio Grande do Sul (6. Câmara Cível*). Apelação Cível nº 70059936971*. Data de Julgamento: 16/12/2014.

ROCK, D.; GRANT, H. Why Diverse Teams Are Smarter. *Harvard Business Review*, 4 nov. 2016. Disponível em: https://hbr.org/2016/11/why-diverse-teams-are-smarter. Acesso em: 27 out. 2020.

RODOTÀ, S. *A vida na sociedade da vigilância*: a privacidade hoje. Tradução: Danilo Doneda e Luciana Cabral Doneda. Rio de Janeiro: Renovar, 2008.

RODOTÀ, S. *El derecho a tener derechos*. Madri: Editorial Trotta, S.A., 2014.

ROSENBLAT, A. *et al.* Discriminating Tastes: Uber's Customer Ratings as Vehicles for Workplace Discrimination. *Policy & Internet,* [*S. l.*], v. 9, n. 3, p. 256–279, 2017. Disponível em: https://doi.org/10.1002/poi3.153. Acesso em: 21 jun. 2021.

RUDIN, C. Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead. *ArXiv*, *[cs, stat]*, [*S. l.*], 2019. Disponível em: http://arxiv.org/abs/1811.10154. Acesso em: 10 ago. 2020.

RUSSON, M. Uber sued by drivers over 'automated robo-firing'. *BBC News*, 26 out. 2020. Disponível em: https://www.bbc.com/news/business-54698858. Acesso em: 1 jun. 2021.

SANDVIG, C. *et al. Auditing Algorithms*: Research Methods for Detecting Discrimination on Internet Platforms. [*S. l.*]: [*s. n*.], 2014.

SÃO PAULO. Assembleia Legislativa do Estado de São Paulo. *CPI – Organizações Sociais da Saúde*. 07 jun. 2018. Disponível em: https://www.al.sp.gov.br/spl/2018/06/Transcricao/1000221166_1000187328_Transcri

cao.pdf. Acesso em: 20 abr. 2021.

SÃO PAULO. GOVERNO DO ESTADO DE SÃO PAULO. Central de Regulação de Oferta de Serviços de Saúde. Disponível em: http://www.cross.saude.sp.gov.br/. Acesso em: 20 abr. 2021.

SAWERS, Paul. Uber drivers union asks EU court to overrule 'robo firing' by algorithm. *Venture Beat*, 26 out. 2020. Disponível em: https://venturebeat.com/2020/10/26/uber-drivers-union-asks-eu-court-to-overrule-robo-firing-by-algorithm/. Acesso em: 10 jan. 2021.

SAXENA, D. *et al.* A Human-Centered Review of Algorithms used within the U.S. Child Welfare System. *In*: CHI '2020, Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. New York, NY, USA: Association for Computing Machinery, 2020. p. 1–15. Disponível em: https://doi.org/10.1145/3313831.3376229. Acesso em: 26 maio 2021.

SECAF, H.; ZANATTA, R. A. F.; NUÑEZ, I. S. O Cadastro Base do Cidadão na mira do Supremo. *Jota*, 9 abr. 2021. Disponível em: https://www.jota.info/opiniao-e-analise/colunas/agenda-da-privacidade-e-da-protecao-de-dados/o-cadastro-base-do-cidadao-na-mira-do-supremo-09042021. Acesso em: 24 abr. 2021.

SELBST, A. D.; POWLES, J. Meaningful Information and the Right to Explanation. *International Data Privacy Law*, v. 7, n. 4, p. 233-242, 2017. Disponível em: https://papers.ssrn.com/abstract=3039125. Acesso em: 27 maio 2020.

SILVA, P. R. Os direitos dos titulares de dados. *In*: MULHOLLAND, Caitlin (Org.). *A LGPD e o novo marco normativo no Brasil*. Porto Alegre: Arquipélago, 2020.

SILVA, T. Reconhecimento facial deve ser banido: veja dez razões. *Tarcízio Silva*, 16 maio 2021. Disponível em: https://tarciziosilva.com.br/blog/reconhecimento-facial-deve-ser-banido-aqui-estao-dez-razoes/. Acesso em: 31 maio 2021.

SINHA, G.; SHAHI, R.; SHANKAR, M. Human Computer Interaction. *In*: INTERNATIONAL CONFERENCE ON EMERGING TRENDS IN ENGINEERING AND TECHNOLOGY, 3., 2010. p. 1–4. Disponível em: https://doi.org/10.1109/ICETET.2010.85. Acesso em: 21 jun. 2021.

SMITH, A. Using Artificial Intelligence and Algorithms. *Federal Trade Commission*, 8 abr. 2020. Disponível em: https://www.ftc.gov/news-events/blogs/business-blog/2020/04/using-artificial-intelligence-algorithms. Acesso em: 23 jul. 2020.

SOUZA, C. A.; PERRONE, C.; MAGRANI, E. O direito à explicação entre a experiência europeia e a sua positivação na LGPD. *In*: MENDES, L. S.; DONEDA, D.; SARLET, I. W.; RODRIGUES JÚNIOR, O. L.; BIONI, B. (Orgs.). *Tratado de Proteção de Dados Pessoais*. Rio de Janeiro: Forense, 2021. p. 243-270.

SPIELKAMP, M. Inspecting Algorithms for Bias. *MIT Technology Review*, 12 jun. 2017. Disponível em: https://www.technologyreview.com/2017/06/12/105804/inspecting-algorithms-for-bias/. Acesso em: 31 maio 2021.

STAN, F. History, motivations, and core themes. *In*: FRANKISH, K.; RAMSEY, W. M. (org.). *The Cambridge Handbook of Artificial Intelligence*. Cambridge: Cambridge University Press, 2014. Disponível em: https://doi.org/10.1017/CBO9781139046855.007. Acesso em: 24 ago. 2020.

STREVENS, M. No understanding without explanation. *Studies in History and Philosophy of Science Part A*. v. 44, n. 3, p. 510–515, 2013. Disponível em: https://doi.org/10.1016/j.shpsa.2012.12.005. Acesso em: 21 jun. 2021.

TAKANO, M. Opening the Black Box of Forensic Algorithms. *Medium*, 3 dez. 2019. Disponível em: https://medium.com/@repmarktakano/opening-the-black-box-of-forensic-algorithms-6194493b9960. Acesso em: 24 jul. 2020.

TENE, Omer. Privacy: The new generations. *International Data Privacy Law*, v. 1, n. 1, p. 15-27, fev. 2011. Disponível em: https://academic.oup.com/idpl/article/1/1/15/759641. Acesso em; 21 jun. 2021.

TEPEDINO, G. *et al.* (Ed.). *Lei Geral de Proteção de Dados Pessoais e suas repercussões no direito brasileiro*. São Paulo: Thomson Reuters, Revista dos Tribunais, 2019.

THALER, R. H.; SUNSTEIN, C. R. *Nudge: improving decisions about health, wealth, and happiness*. New York: Penguin Books, 2009.

TOMCZYK, L. *et al*. Digital Divide in Latin America and Europe: Main Characteristics in Selected Countries. *In*: IBERIAN CONFERENCE ON INFORMATION SYSTEMS AND TECHNOLOGIES (CISTI), 14., 2019. Disponível em: https://doi.org/10.23919/CISTI.2019.8760821. Acesso em: 21 jun. 2021.

TREASURY BOARD OF CANADA. *Directive on Privacy Impact Assessment*. 2010. Disponível em: https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=18308. Acesso em: 11 abr. 2021.

UNIÃO EUROPEIA. *Regulamento Geral sobre a Proteção de Dados* — Considerando 71. Disponível em: https://gdpr-text.com/pt/read/recital-71/. Acesso em: 21 jun. 2021.

UNIÃO EUROPEIA. *WP 173* — Opinion 3/2010 on the principle of accountability. [*S. l.*]: WP29, 2010. Disponível em: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2010/wp173_en.pdf. Acesso em: 21 jun. 2021.

UNIÃO EUROPEIA. *WP 260rev.01* — Guidelines on Transparency under Regulation 2016/679. [*S. l.*]: WP29, 2018. Disponível em: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=622227. Acesso em; 21 jun. 2021.

UNIÃO EUROPEIA. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing*

Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance). Official Journal of the European Union, 4 maio 2016. Disponível em: https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng. Acesso em: 27 ago. 2021.

UNIÃO EUROPEIA. *Tratado sobre o Funcionamento da União Europeia*. 6 jul. 2016. Disponível em: https://eur-lex.europa.eu/resource.html?uri=cellar:9e8d52e1-2c70-11e6-b497 01aa75ed71a1.0019.01/DOC_3&format=PDF. Acesso em: 25 nov. 2020.

UNIÃO EUROPEIA. *COM (2018) 237 — Artificial Intelligence for Europe*. 2018. Disponível em: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A237%3AFIN. Acesso em: 13 jul. 2020.

UNIÃO EUROPEIA. *COM (2018) 785 — Coordinated Plan for Artificial Intelligence*, 2018. Disponível em: https://ec.europa.eu/knowledge4policy/publication/coordinated-plan-artificial-intelligence-com2018-795-final_en#:~:text=7%20December%202018—,Coordinated%20Plan%20on%20Artificial%20Intelligence%20. Acesso em 13 jul. 2020.

UNIÃO EUROPEIA. *COM (2019) 168 — Building Trust in Human-Centric Artificial Intelligence*. 2019. Disponível em: https://ec.europa.eu/transparency/regdoc/rep/1/2019/EN/COM-2019-168-F1-EN-MAIN-PART-1.PDF. Acesso em: 13 jul. 2020.

UNIÃO EUROPEIA. *Ethics Guidelines for Trustworthy Artificial Intelligence*. European Commission, 2019. Disponível em: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai. Acesso em: 13 jul. 2018.

UNIÃO EUROPEIA. *COM (2020) 65 — White Paper on Artificial Intelligence: a European approach to excellence and trust*. Disponível em: https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en. Acesso em: 01 ago. 2020.

UNIÃO EUROPEIA. Tribunal de Justiça da União Europeia (2. Câmara). *Caso C-434/16, Peter Nowak v. Data Prot. Comm'r,* Data de Julgamento: 20/12/2017. Disponível em: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A62016CJ0434. Acesso em: 21 jun. 2021.

UNIÃO EUROPEIA. Tribunal de Justiça da União Europeia (3. Câmara). *Casos C-141/12 & 372/12, YS, M e S v. Minister voor Immigratie, Integratie en Asiel*, Data de Julgamento: 17/06/2014. Disponível em: https://curia.europa.eu/juris/liste.jsf?num=C-141/12&language=en. Acesso em: 21 jun. 2021.

VAN ALSENOY, B. Liability under EU Data Protection Law: From Directive 95/46 to the General Data Protection Regulation. *JIPITEC*, [*S. l.*], v. 7, n. 3, 2017. Disponível em: http://www.jipitec.eu/issues/jipitec-7-3-2016/4506. Acesso em: 21 jun; 2021.

VEALE, M. *Governing Machine Learning that Matters*. 2019. Doctoral Thesis (PhD in Science, Technology, Engineering and Public Policy) — University College London, London, 2019. Disponível em:

https://discovery.ucl.ac.uk/id/eprint/10078626/1/thesis_final_corrected_mveale.pdf. Acesso em 15 dez. 2020.

VELOSO, A. C; CARDOSO, L.; BRÊTAS, P. Jogo dos 7 erros: auxílio de R$ 600 é negado a quem tem requisites e concedido a quem não precisa. *O Globo*, 5 jun. 2020. Disponível em: https://oglobo.globo.com/economia/jogo-dos-7-erros-auxilio-de-600-negado-quem-tem-requisitos-concedido-quem-nao-precisa-1-24464513. Acesso em: 20 abr. 2021.

VILLANI, C. *For a Meaningful Artificial Intelligence:* Towards a French and European Strategy. [*S. l.*]: European Commission, 2018. Disponível em: https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf. Acesso em: 28 jun. 2021.

WACHTER, S.; MITTELSTADT, B. A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI. *LawArXiv*, 12 out. 2018. Disponível em: https://doi.org/10.31228/osf.io/mu2kf. Acesso em: 27 maio 2020.

WACHTER, S.; MITTELSTADT, B. A Right to Reasonable Inferences: Re-thinking Data Protection Law in the Age of Big Data and AI. *Columbia Business Law Review*, 2019, v. 2, 5 out. 2018. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3248829. Acesso em: 16 abr. 2021.

WACHTER, S.; MITTELSTADT, B.; FLORIDI, L. *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation*. Rochester, NY: Social Science Research Network, 2016. Disponível em: https://papers.ssrn.com/abstract=2903469. Acesso em: 27 maio. 2020.

WACHTER, S.; MITTELSTADT, B.; FLORIDI, L. Why a Right to Explanation of Automated Decision Making Does Exist in the General Data Protection Regulation. *In: International Data Privacy Law*, vol. 7, n. 2, maio 2017, p. 76–99. Disponível em: https://academic.oup.com/idpl/article/7/2/76/3860948. Acesso em: 6 nov. 2020.

WACHTER, S.; MITTELSTADT, B.; RUSSELL, C. Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR. *SSRN Electronic Journal*, v. 31, n. 2, 2017. Disponível em: https://www.ssrn.com/abstract=3063289. Acesso em: 27 maio. 2020.

WALKER, S. M. *The future of human rights impact assessments of trade agreements*. [*S. l.*]: [*s. n*.], 2009.

WARREN, S. D.; BRANDEIS, L. D. The Right to Privacy. *Harvard Law Review*, v. 4, n. 5, 1890.

WATSON, D. S.; FLORIDI, L. The explanation game: a formal framework for interpretable machine learning. *Synthese*, 2020. Disponível em: https://doi.org/10.1007/s11229-020-02629-9. Acesso em: 8 jun. 2020.

WIMMER, M. Autoridades de proteção de dados pessoais no mundo: fundamentos e

evolução na experiência comparada. *In:* PALHARES, Felipe (Coords.). *Temas atuais de proteção de dados*. São Paulo: Thomson Reuters, 2020.

WIMMER, M. Os desafios do *enforcement* na LGPD: fiscalização, aplicação de sanções administrativas e coordenação intergovernamental. *In:* MENDES, L. S.; DONEDA, D.; SARLET, I. W.; RODRIGUES JR., O. L. (Coords.). *Tratado de Proteção de Dados Pessoais*. Rio de Janeiro: Forense, 2021.

WORLD TRADE ORGANIZATION. *Trade-Related Aspects of Intellectual Property Rights*. 15. abr. 1994. Disponível em: https://wipolex.wipo.int/en/text/305907. Acesso em: 26 abr. 2021.

WRIGHT, D.; MORDINI, E. Privacy and Ethical Impact Assessment. *In*: WRIGHT, D.; DE HERT, P. (org.). *Privacy Impact Assessment.* Dordrecht: Springer Netherlands, 2012. p. 397–418. Disponível em: https://doi.org/10.1007/978-94-007-2543-0_19. Acesso em: 24 abr. 2021.

ZANATTA, R. A. F. *Pontuação de Crédito e Direitos dos Consumidores*: o desafio brasileiro. São Paulo: Idec, 2017.

ZANATTA, R. A. F. A tutela coletiva na proteção de dados pessoais. *Revista do Advogado*, n. 144, nov. 2019. Disponível em: https://aplicacao.aasp.org.br/aasp/servicos/revista_advogado/paginaveis/144/200/index.html. Acesso em: 20 maio 2021.

ZANATTA, R. A. F. Agentes De Tratamento De Dados, Atribuições e Diálogo Com O Código De Defesa Do Consumidor. *Revista dos Tribunais*, [*S. l.*], n. 1009, supl. Caderno Especial, nov. 2019.

ZANATTA, R. A. F. Tutela coletiva e coletivização da proteção de dados pessoais. *In:* PALHARES, F. (Org.). *Temas Atuais de Proteção de Dados*. São Paulo: Thomson Reuters, 2020.

ZANATTA, R. A. F.; PAULA, P. C. B. de; KIRA, B. *Economias do Compartilhamento e o Direito*. 1. ed. São Paulo: Juruá Editora, 2017.