

INTELIGÊNCIA ARTIFICIAL, DISCRIMINAÇÃO ALGORÍTMICA E O DEVER DE DILIGÊNCIA DOS ADMINISTRADORES DAS SOCIEDADES ANÔNIMAS

Camila Trindade Galvão



Editora Fundação Fênix



O livro aborda o fenômeno da discriminação algorítmica perpetrada por sistemas de inteligência artificial, explorando as suas repercussões na interpretação do dever de diligência dos administradores de empresas, conforme estipulado no artigo 153 da Lei 6.404/76. Fruto da dissertação de mestrado da autora, o livro oferece conclusões que destacam a necessidade de uma interpretação mais específica desse dever em relação à inteligência artificial. Ao identificar os deveres que os administradores devem cumprir para que suas decisões negociais sejam consideradas diligentes, a autora busca oferecer orientações sobre as ações que os administradores de empresas que utilizam inteligência artificial devem adotar para evitar ou mitigar a discriminação algorítmica. A pesquisa conclui que, junto com a IA, surge a necessidade de implementar programas de compliance e estruturas de governança capazes de enfrentar a discriminação algorítmica. Uma obra essencial para profissionais do direito, administradores e pesquisadores que desejam compreender os desafios legais e éticos da inteligência artificial nos negócios da atualidade.



Editora Fundação Fênix



**INTELIGÊNCIA ARTIFICIAL, DISCRIMINAÇÃO ALGORÍTMICA E O DEVER DE
DILIGÊNCIA DOS ADMINISTRADORES DAS SOCIEDADES ANÔNIMAS**

Série Direito

Conselho Editorial

Editor

Ingo Wolfgang Sarlet

Conselho Científico – PPG Direito PUCRS

Gilberto Stürmer – Ingo Wolfgang Sarlet

Marco Felix Jobim – Paulo Antonio Caliendo Velloso da Silveira

Regina Linden Ruaro – Ricardo Lupion Garcia

Conselho Editorial Nacional

Adalberto de Souza Pasqualotto – PUCRS

Amanda Costa Thomé Travincas – Centro Universitário UNDB

Ana Elisa Liberatore Silva Bechara – USP

Ana Maria DÁvila Lopes – UNIFOR

Ana Paula Gonçalves Pereira de Barcellos – UERJ

Angélica Luciá Carlini – UNIP

Augusto Jaeger Júnior – UFRGS

Carlos Bolonha – UFRJ

Claudia Mansani Queda de Toledo – Centro Universitário Toledo de Ensino de Bauru

Cláudia Lima Marques – UFRGS

Clara Iglesias Keller – WZB Berlin Social Sciences Center e Instituto Brasileiro de Ensino

Desenvolvimento e Pesquisa – IDP

Danielle Pamplona – PUCRS

Daniel Antônio de Moraes Sarmiento – UERJ

Daniel Wunder Hachem – PUCPR e UFPR

Daniel Mitidiero – UFRGS

Denise Pires Fincato – PUCRS

Draiton Gonzaga de Souza – PUCRS

Eugênio Facchini Neto – PUCRS

Elda Coelho de Azevedo Bussinguer – UniRio

Fabio Siebeneichler de Andrade – PUCRS

Fabiano Menke – UFRGS

Flavia Cristina Piovesan – PUC-SP

Gabriel de Jesus Tedesco Wedy – UNISINOS

Gabrielle Bezerra Sales Sarlet – PUCRS

Germano André Doederlein Schwartz – UNIRITTER

Gilmar Ferreira Mendes – Ministro do STF, Professor Titular do IDP e Professor aposentado da UNB

Gisele Cittadino – PUC-Rio

Gina Vidal Marcilio Pompeu – UNIFOR

Giovani Agostini Saavedra – Universidade Presbiteriana Mackenzie – SP

Guilherme Camargo Massaú – UFPel

Gustavo Osna – PUCRS

Hermes Zaneti Jr

Hermilio Pereira dos Santos Filho – PUCRS
Ivar Alberto Martins Hartmann – FGV Direito Rio
Jane Reis Gonçalves Pereira – UERJ
Juliana Neuenschwander Magalhães - UFRJ
Laura Schertel Mendes
Lilian Rose Lemos Rocha – Uniceub
Luís Alberto Reichelt – PUCRS
Luís Roberto Barroso – Ministro do STF, Professor Titular da UERJ, UNICEUB, Sênior Fellow na Harvard Kennedy School
Miriam Wimmer - IDP - Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa
Mônia Clarissa Hennig Leal – UNISC
Otavio Luiz Rodrigues Jr – USP
Patryck de Araújo Ayala – UFMT
Paulo Ricardo Schier - Unibrasil
Phillip Gil França - UNIVEL – PR
Richard Pae Kim – UNISA
Teresa Arruda Alvim – PUC-SP
Thadeu Weber – PUCRS

Conselho Editorial Internacional

Alexandra dos Santos Aragão – Universidade de Coimbra
Alvaro Avelino Sanchez Bravo – Universidade de Sevilha
Catarina Isabel Tomaz Santos Botelho – Universidade Católica Portuguesa
Carlos Blanco de Morais – Universidade de Lisboa
Clara Iglesias Keller – WZB Berlin Social Sciences Center e Instituto Brasileiro de Ensino
Desenvolvimento e Pesquisa – IDP
Cristina Maria de Gouveia Caldeira – Universidade Europeia
César Landa Arroyo – PUC de Lima, Peru
Elena Cecilia Alvites Alvites – Pontifícia Universidade Católica do Peru
Elena Alvites Alvites - PUCP
Francisco Pereira Coutinho – Universidade NOVA de Lisboa
Francisco Ballaguer Callejón – Universidade de Granada - Espanha
Fernando Fita Ortega - Universidade de Valência
Giuseppe Ludovico - Universidade de Milão
Gonzalo Aguilar Cavallo – Universidade de Talca
Jorge Pereira da Silva – Universidade Católica Portuguesa
José João Abrantes – Universidade NOVA de Lisboa
José Maria Porrás Ramirez – Universidade de Granada – Espanha
Manuel A Carneiro da Frada – Universidade do Porto
Paulo Mota Pinto – Universidade de Coimbra
Pedro Paulino Grandez Castro – Pontificia Universidad Católica del Peru
Richard Pae Kim – Professor do Curso de Mestrado em Direito Médico da UNSA
Víctor Bazán – Universidade Católica de Cuyo

Camila Trindade Galvão

**INTELIGÊNCIA ARTIFICIAL, DISCRIMINAÇÃO ALGORÍTMICA E O DEVER DE
DILIGÊNCIA DOS ADMINISTRADORES DAS SOCIEDADES ANÔNIMAS**



Editora Fundação Fênix

Porto Alegre, 2023

Direção editorial: Ingo Wolfgang Sarlet
Diagramação: Editora Fundação Fênix
Capa: Editora Fundação Fênix

O padrão ortográfico, o sistema de citações, as referências bibliográficas, o conteúdo e a revisão de cada capítulo são de inteira responsabilidade de seu respectivo autor.

Todas as obras publicadas pela Editora Fundação Fênix estão sob os direitos da Creative Commons 4.0 –
http://creativecommons.org/licenses/by/4.0/deed.pt_BR



Série Direito – 81

Catálogo na Fonte

G182i Galvão, Camila Trindade
Inteligência artificial, discriminação algorítmica e o dever de diligência dos administradores das sociedades anônimas [recurso eletrônico] / Camila Trindade Galvão. – Porto Alegre : Editora Fundação Fênix, 2023.
187 f. (Série Direito ; 81)

Dissertação (Mestrado) – Pontifícia Universidade Católica do Rio Grande do Sul, Programa de Pós-Graduação em Direito, 2023.
Orientador: Prof. Dr. Ricardo Lupion Garcia.
Disponível em: <<http://www.fundarfenix.com.br>>
ISBN 978-65-5460-061-3
DOI <https://doi.org/10.36592/9786554600613>

1. Inteligência artificial. 2. Algoritmos. 3. Direito administrativo.
4. Direito. I. Garcia, Ricardo Lupion (orient.). II. Título

CDD: 340

Responsável pela catalogação: Lidiane Corrêa Souza Morschel CRB10/1721

AGRADECIMENTOS

Ao meu amor, que no último ano foi promovido de noivo a marido, agradeço a parceria em me acompanhar nesta loucura que foi o mestrado, me incentivando a ir longe para realizar meus sonhos! Agradeço também por compreender tantas ausências e tantos nervosismos, sendo o suporte mais do que necessário para que eu pudesse superar todos os meus medos, estando sempre disposto a ajudar com muita dedicação, calma e paciência. Te amo!

Agradeço também aos meus pais, avós, sogros, manas, cunhados, sobrinhos, demais familiares e amigas, que tanto vibraram e compreenderam a complexidade e a importância deste momento na minha vida. Eu não sou nada sem vocês!

A minha gratidão pelo meu super orientador também precisa ser registrada! O professor Ricardo Lupion foi muito mais do que um orientador. Ele foi um guia, um protetor, um líder, um amigo, um consolo, um parceiro de trabalhar muito e de solucionar desafios, um construtor de pontes que acreditou em mim e me abriu muitas oportunidades, um excelente motivador. Foi firme e exigente quando precisava e foi descontraído e acolhedor quando cabia. Sempre amparou meus medos e anseios, disponível para me aconselhar até sobre situações acadêmicas alheias à minha pesquisa. Um professor que consegue sustentar todo requinte dos inúmeros títulos acadêmicos que possui sem perder a praticidade, a objetividade, a organização e a humildade. Preocupadíssimo com a experiência do aluno, valorizando o melhor de cada um, sempre aberto a diálogos e a colaborações. Obrigada por esta convivência. Dentre todos os altos e baixos da minha trajetória acadêmica, poder conviver, observar e aprender com um verdadeiro mestre foi a melhor experiência que eu posso levar desse mestrado (e olha que aprendi muito e fiz grandes amigos). Se eu conseguir ser 10% disso para os meus futuros alunos, já me sinto vitoriosa! SUPER obrigada!

Agradeço também à professora Regina Linden Ruaro, por ser uma inspiração tão grande quanto a sua generosidade e naturalidade ao compartilhar o que sabe. A admiração, que já existia, só cresceu!

À Fernanda Linden Ruaro Peringer, agradeço por me introduzir, me encorajar e me indicar os caminhos para esta jornada acadêmica.

Aos meus colegas de mestrado, especialmente os "Giraffers", agradeço por dividirem comigo as dores, as incertezas e as conquistas desta etapa que nos exigiu tanta energia. O nosso grupo de apoio foi fundamental para este processo.

À Rosana Arruda Pagnussat, por toda dedicação, gentileza e paciência para organizar este trabalho dentro das normas da ABNT.

E por último, mas não menos importante, agradeço ao time do Galvão Advogados, em especial ao meu parceiro de fé, Fabio Portal, à Virginia Galvão, ao Breno Galvão, ao Rafael Pernigotti, à Priscilla Pernigotti e à nossa equipe de gestores. Seria impossível dar conta de todas as demandas do mestrado se não houvesse vocês para auxiliar a administrar as minhas ausências e me representar sempre que necessário. Sobrevivemos!

Lista de Figuras

Figura 1 – Exemplos de rotulação	37
Figura 2 – Inteligência artificial vs humana	40
Figura 3 – Avisos dados antes de utilizar o Chat GPT	66

Lista de Quadros

Quadro 1 – Tipos de generalizações	50
Quadro 2 – Discriminações algorítmicas estatísticas	51
Quadro 3 – A previsão falha de forma diferente para réus negros	61
Quadro 4 – Quadro-resumo discriminação algorítmica conforme Mendes e Mattiuzzo	74
Quadro 5 – Quadro-resumo discriminação algorítmica conforme Doneda <i>et al.</i>	75

SUMÁRIO

PREFÁCIO

Ricardo Lupion 15

1. INTRODUÇÃO 19

2. INTELIGÊNCIA ARTIFICIAL E UMA NOVA FORMA DE DISCRIMINAÇÃO 25

2.1 A expansão da Inteligência Artificial 25

2.2 Sobre a Inteligência Artificial (IA) 27

2.3 Algoritmos 31

2.4 Aprendizagem de máquina 33

2.5 Riscos e aspectos jurídicos da IA 41

2.6 Discriminação algorítmica 47

2.7 Como ocorre a discriminação algorítmica 58

2.8 Possíveis causas 68

2.9 Conclusões parciais 81

3. DO DEVER DE DILIGÊNCIA 83

3.1 Exame da doutrina 83

3.2 Dever de se informar 97

3.3 Dever de se qualificar 101

3.4 Dever de bem administrar 103

3.5 Dever de investigar 104

3.6 Dever de vigiar/vigilância 108

3.7 Dever de supervisão/monitoramento 111

3.8 Dever de intervir 112

3.9 Dever de não praticar erros graves 113

3.10 Dever de participação/assiduidade 114

3.11 Outros deveres envolvidos 114

3.12 *Business Judgment Rule* 119

3.13 A Comissão de Valores Mobiliários (CVM) e o dever de diligência 122

3.14 Conclusões parciais 127

4. A DISCRIMINAÇÃO ALGORÍTMICA E O DEVER DE DILIGÊNCIA	131
4.1 <i>Compliance</i>	136
4.2 Entendendo as fontes de vieses e suas possíveis soluções: o <i>compliance</i> pode evitar ou mitigar a discriminação algorítmica?	144
4.3 <i>Compliance</i> algorítmico e o dever de diligência	154
4.4 Conclusões parciais	169
5. CONCLUSÃO FINAL	171
REFERÊNCIAS	177

PREFÁCIO

O presente livro é resultado da dissertação apresentada para a obtenção do título de Mestre em Direito no Programa de Pós-Graduação *stricto sensu* da Pontifícia Universidade Católica do Rio Grande do Sul.

Tive a satisfação de participar da pesquisa como orientador e de presidir a banca examinadora da dissertação. O trabalho foi aprovado com grau máximo com recomendação para publicação, por banca examinadora do mais alto nível, composta pelo Prof. Dr. Oksandro Osdival Gonçalves, da Escola de Direito da Pontifícia Universidade Católica do Paraná, (PUCPR), pelo Prof. Dr. Cristiano Colombo, da Escola de Direito da Universidade do Vale do Rio dos Sinos, (UNISINOS) e pela Prof^a. Dr^a. Regina Linden Ruaro, da Escola de Direito da Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), realizada no dia 14/03/2023.

A obra é de especial relevância para os pesquisadores, profissionais e estudiosos da inteligência artificial e a sua conexão com o direito societário, porque o uso e aplicação dessa tecnologia normalmente está vinculado à atividade de uma empresa.

A dissertação do mestrado, cujo texto resultou no presente livro, é resultado de uma rigorosa pesquisa sobre tema de grande relevância: os riscos e desafios que o uso da inteligência artificial em larga escala vem apresentando, com análise de seus reflexos jurídicos. Esta obra surge quando vivemos a quarta revolução industrial, era em que a tecnologia é cada vez mais utilizada e participa ostensivamente de nossas vidas, ainda que muitas vezes ocorra de forma despercebida.

As decisões tomadas por algoritmos de inteligência artificial, abordadas no livro, se tornam cada vez mais frequentes, proporcionando inúmeras facilidades e avanços, igualmente proporcionais aos riscos que podem ocasionar, muitos inerentes à sua própria configuração. Os algoritmos podem influenciar comportamentos, moldar destinos e afetar a sociedade em tal escala que até ultrapassa a atual compreensão humana.

Assim, a autora descreve uma importante preocupação envolvendo o assunto, a discriminação algorítmica, e elenca as possíveis causas e consequências desta nova forma de propagar preconceitos e segregar grupos de pessoas, o que pode ocorrer porque muitas empresas ainda não estão habituadas a lidar com este tema.

Na sequência, a autora, em sua pesquisa, além de descrever os atributos do dever de diligência dos administradores de uma sociedade anônima, com base em renomada doutrina societária, também avaliou como a adoção de novas práticas e ferramentas no ambiente corporativo (como o uso de IA, por exemplo) pode, ao lado dos atributos tradicionais, integrar os requisitos do dever de diligência no processo decisório do administrador pois, como antes referido, o uso da inteligência artificial sempre envolverá a atividade empresarial.

Mas a pesquisa elaborada foi além das características comuns de uma dissertação de mestrado: foi propositiva porque não se limitou a descrever os aspectos relacionados à inteligência artificial e como o uso dessa nova tecnologia pode impactar os atributos do dever de diligência dos administradores de uma sociedade anônima (capítulos primeiro e segundo), foi além: no terceiro e último capítulo, procurou identificar como o dever de diligência pode ser concretizado e reconhecido nestas situações, constatando-se que os instrumentos adequados para instrumentalizá-lo podem ser a criação de estruturas de governança, especialmente o compliance com o finalidade de mitigar discriminações algorítmicas – o que, por via reflexa, também auxilia no cumprimento de outro dever: o de legalidade, ou seja, o dever de não desrespeitar normas legais.

Tive a felicidade de participar dessa trajetória na companhia da agora mestre em direito, Camila Galvão. A Camila foi bolsista CAPES e participou ativamente das atividades acadêmicas durante a sua vinculação ao PPGD/PUCRS, sempre demonstrando preocupação e cuidado extremos com o seu projeto de pesquisa, com o desenvolvimento da sua dissertação de mestrado e elevado comprometimento com todas as atividades do PPGD/PUCRS.

Por tudo isso, comecei como orientador e agora sou amigo e admirador dessa jovem e talentosa acadêmica e advogada, a quem desejo saúde e sucesso!

Estimo que esta obra auxilie e inspire estudos futuros, cumprindo a missão de colaborar significativamente para o debate sobre a inteligência artificial, seus desafios, oportunidades e, sobretudo, sua regulação.

Desejo a todos uma excelente leitura!

Porto Alegre, agosto de 2023.

Ricardo Lupion.

Agente de Inovação da Escola de Direito da PUCRS.

Professor do Programa de Pós-Graduação em Direito da PUCRS.

1. INTRODUÇÃO

Com o aumento exponencial da tecnologia na vida de todos, surgiram muitos riscos imprevistos, ocasionando ofensas a direitos e garantias que jamais seriam aceitas se perpetradas por cidadãos humanos e que estão associadas ao uso massivo de ferramentas tecnológicas. Além disso, a velocidade dos avanços é superior ao tempo de compreensão dos problemas, bem como da adaptação das normas jurídicas para freá-los e incentivar os benefícios, de forma que há um grande desafio a ser solucionado.

Após muitas pesquisas que serão mencionadas adiante, constatou-se que uma das tecnologias responsáveis pelos rápidos e, por vezes, "descontrolados" avanços, é a inteligência artificial, presente em computadores, *smartphones*, televisores, relógios, veículos, livros digitais, equipamentos médicos, indústrias, entre outros diversos dispositivos.

Analisando de forma geral, a inteligência artificial proporciona um sem-número de novas possibilidades e facilidades, que beneficiam muito a sociedade. Contudo, ao tentar reproduzir a inteligência humana, reproduz-se, também, suas limitações, ensejando reflexos negativos.

Um destes reflexos negativos é a possibilidade de algumas aplicações afrontarem o princípio da igualdade, discriminando indivíduos de forma silenciosa e em grande escala, ainda que contra a vontade de seus idealizadores e programadores. Isso porque, em muitos casos, a inteligência artificial se adapta a partir dos dados produzidos pela própria sociedade, sem um juízo crítico de valor, e pode acabar reproduzindo preconceitos históricos e, infelizmente, presentes em registros digitais.

Eis que assim nasceu o problema desta pesquisa: criou-se uma ferramenta inovadora, cujo conhecimento e domínio técnico está concentrado em uma pequena parcela de pessoas e empresas, que traz inúmeros benefícios, mas também pode ocasionar resultados indesejados até mesmo por quem a planejou. Estes resultados podem representar ilicitudes e inconstitucionalidades, ainda que sejam gerados a partir de circunstâncias totalmente avessas às empresas desenvolvedoras. Por sua vez, tais empresas se encontram sujeitas a sérios riscos e prejuízos, eis que, mesmo

utilizando-se dos melhores recursos técnicos existentes no mercado, as discriminações podem ocorrer de forma não intencional.

Então, questiona-se:

a) O que o ordenamento jurídico espera das empresas que utilizam inteligência artificial e de seus respectivos gestores?

b) Que cautelas podem ou devem ser tomadas para evitar que as decisões proferidas por suas inteligências artificiais representem um desrespeito à ordem jurídica?

c) E, em caso de não ser possível evitar todos os danos, o que se deve fazer para evitar que as condutas dos administradores representem omissão, descuido ou falta de cautelas por parte das empresas responsáveis, sobretudo diante da existência de um dever que exige diligência dos seus administradores na condução dos negócios, consoante a previsão do artigo 153 da Lei das Sociedades Anônimas (Lei 6.404/76)?¹

d) A interpretação atualizada do dever de diligência impõe alguma obrigação aos administradores das empresas que usam ou desenvolvem inteligência artificial em suas rotinas?

A hipótese a ser verificada é se o dever de diligência exigido dos administradores pode ser interpretado de forma a obrigá-los a adotar esta mesma diligência nas tomadas de decisão que envolvem o uso de inteligência artificial, especialmente aquelas que são utilizadas em processos decisórios que ofertam (ou negam) serviços, bens e direitos a indivíduos e podem ser discriminatórias.

Assim, para sanar as dúvidas anteriormente mencionadas, o estudo buscará aprofundar o entendimento sobre o atual estado da arte da inteligência artificial e, mais especificamente, o que significa a discriminação algorítmica e como ela ocorre, o que se fará no primeiro capítulo. Neste item, será relatada uma pesquisa bibliográfica, analisando-se doutrina jurídica especializada, bem como pesquisas da

¹ LEI nº. 6.404, DE 15 DE DEZEMBRO DE 1976. (Dispõe sobre as Sociedades por Ações.): Art. 153. O administrador da companhia deve empregar, no exercício de suas funções, o cuidado e diligência que todo homem ativo e probo costuma empregar na administração dos seus próprios negócios. BRASIL. **Lei n. 6.404, de 15 de dezembro de 1976.** Dispõe sobre as sociedades por ações. Brasília: Casa Civil, 1976. Disponível em: https://www.planalto.gov.br/ccivil_03/Leis/L6404consol.htm. Acesso em: 15 abr. 2022.

própria área computacional, documentadas em livros, artigos, vídeos e notícias. O principal objetivo deste capítulo será compreender quais fatores podem contribuir para a ocorrência da discriminação preconceituosa, para, após, avaliar as medidas possíveis de serem exigidas, do ponto de vista fático e jurídico.

Deste modo, o texto iniciará por uma breve contextualização sobre a expansão da inteligência artificial, situando o leitor do seu papel central na chamada Quarta Revolução Industrial, momento histórico vivido atualmente, conforme apontado por autores.

Na sequência, a pesquisa se dedicará a estabelecer um acordo semântico com o leitor, esclarecendo qual é o conceito de inteligência artificial adotado na investigação, tendo em vista a ausência de uma definição uníssona sobre esta tecnologia, dadas as diversas abordagens computacionais que a compõem.

Igualmente, o estudo sistematizará os conceitos de algoritmos e aprendizagem de máquina (*machine learning*), eis que relevantes para o entendimento do que é a discriminação algorítmica e suas causas. Entende-se que sem a compreensão do funcionamento da tecnologia investigada, haverá dificuldade de se concluir o que pode ser exigido dos atores envolvidos pelo ordenamento jurídico.

Já ultrapassada a compreensão dos conceitos básicos envolvidos no tema, serão debatidos os diferentes tipos de discriminação algorítmica e suas possíveis causas, além dos direitos violados por eventuais aplicações de inteligência artificial que incorrerem em tais lesões.

Ato seguinte, o segundo capítulo buscará examinar qual é a diligência esperada de um administrador do século XXI, sobretudo aquele que atua no ano de 2023, considerando o fato de que desde a edição da Lei 6.404/76, com o art. 153, quase cinco décadas se passaram e o acesso a informações e dados, que podem subsidiar as tomadas de decisão negociais, alterou-se muito. Assim, a pesquisa trará a evolução da interpretação que a doutrina do direito societário dá ao art. 153 da Lei 6.404/76, bem como da interpretação que é dada pela Comissão de Valores Mobiliários (o que se fará com base em exemplos emblemáticos, dado o grande volume de precedentes que abordam o dever de diligência) para, a partir deste

estudo, elencar quais são os deveres que o administrador de uma empresa deve adotar em qualquer decisão negocial.

A seguir, será realizado um exame em ordem cronológica de dezessete doutrinadores que lecionaram sobre o tema, avaliando as diferentes interpretações dadas ao dever de diligência com o passar do tempo. Após, a pesquisa consolidará todos os deveres e subdeveres mencionados pelos autores pesquisados e relatará em ordem de maior quantidade de menções as suas definições, abordando-se os seguintes: dever de se informar, dever de se qualificar, dever de bem administrar, dever de investigar, dever de vigiar, dever de supervisionar, dever de intervir, dever de não praticar erros graves, dever de participar, bem como outros subdeveres que estão sendo mencionados em pesquisas mais recentes, em especial o dever de adotar estruturas de governança e *compliance*.

Ainda no segundo capítulo, será abordada a relevância da observância do dever de diligência para aplicação da *Business Judgment Rule*, bem como a interpretação que a Comissão de Valores Mobiliários dá para os processos decisórios que conduziram ações práticas dos administradores de empresas.

Ao final, no terceiro capítulo, buscar-se-á unir as duas investigações para interpretar quais são as ações que os administradores de empresas que desenvolvem ou utilizam inteligência artificial em suas decisões e processos negociais devem adotar para evitar e/ou mitigar ilicitudes ou inconstitucionalidades que decorram de suas aplicações de inteligência artificial ou, ainda que ocorram, como evidenciar que tais administradores foram zelosos na adoção de todas as medidas possíveis para evitá-las – tudo isso à luz das mais recentes interpretações a respeito do comando prático contido no artigo 153 da Lei 6.404/76, que vislumbram os programas de conformidade (*compliance*) como a metodologia de concretização do dever de diligência.

O terceiro capítulo iniciará com uma reflexão sobre a interpretação conjunta a respeito do dever de diligência aplicado em processos decisórios que definem a utilização (ou não) de inteligência artificial pelas empresas, bem como os fluxos necessários antes, durante e após a sua implementação. Neste sentido, a pesquisa dedicará um item para esclarecer o que é *compliance*, de modo a fornecer ao leitor o embasamento necessário a formar o juízo crítico que avaliará se estes

procedimentos de conformidade podem ser úteis para utilização de uma inteligência artificial mais segura.

Na sequência, a pesquisa seguirá pela compreensão das fontes de vieses algorítmicos, o que fará a partir de pesquisas da área computacional, com o fito de entender se o *compliance* pode ser útil para evitar ou mitigar os potenciais focos de discriminação algorítmica. Após esta análise, a pesquisa traçará os mecanismos citados pela doutrina como essenciais a um *compliance* algorítmico, sem a intenção de esgotar a matéria.

Importa destacar que pesquisar qualquer tema envolvendo tecnologia torna o trabalho muito dinâmico, ao passo em que inúmeras mudanças e novidades surgem em um curto espaço de tempo. Assim, para não tornar esta análise inesgotável, delimitou-se a um problema muito específico (a discriminação algorítmica), no intuito de tentar aprofundá-lo e, efetivamente, propor algum tipo de solução, a ser extraída a partir da interpretação do conteúdo do dever de diligência.

O método utilizado foi o método dedutivo, eis que, a partir de duas premissas gerais, buscar-se-á uma interpretação mais específica e conjugada do dever de diligência aplicado à inteligência artificial.

Desde o início desta investigação (em março de 2021) até a conclusão do texto (janeiro de 2023) novas tecnologias surgiram, diversos estudos foram publicados e houve um avanço muito rápido – até no aspecto legal, tendo em vista o desenvolvimento do texto que possivelmente irá representar o Marco Legal da Inteligência Artificial, elaborado por grandes juristas brasileiros, no intuito de regular a inteligência artificial no Brasil. A aprovação deste marco pode tornar o Brasil um dos pioneiros na regulação deste tipo de tecnologia, servindo de referência para diversos países com condições semelhantes às do Brasil, especialmente porque dedicou seções específicas para abordar os assuntos desta pesquisa (discriminação algorítmica e governança), reforçando a importância de promover a igualdade no ambiente tecnológico de forma ativa.

Além disso, o resultado deste trabalho poderá auxiliar gestores que buscam obter relativa segurança jurídica no uso e desenvolvimento de ferramentas tecnológicas, eis que, para além dos muros acadêmicos, há empresas de todos os portes que ainda carecem de diretrizes práticas e objetivas para produzir inovação

24 | Inteligência artificial, discriminação algorítmica...

em conformidade com a legislação e sem causar danos a terceiros. De outro lado, há advogados que buscam guarnecer as empresas das melhores orientações e precisam estar em constante estudo e atualização sobre o funcionamento das novas tecnologias, para seu correto enquadramento dentro das normas aplicáveis, situações que justificam a pertinência do conteúdo a ser abordado.

2. INTELIGÊNCIA ARTIFICIAL E UMA NOVA FORMA DE DISCRIMINAÇÃO

2.1 A expansão da Inteligência Artificial

A organização social, econômica e política está mudando e isso está ocorrendo em larga escala e com abrangência global. A cada dia, novos arranjos socioeconômicos surgem a partir de invenções até então existentes apenas na imaginação.

A afirmação acima não é uma novidade, afinal, a sociedade sempre se modificou. Todavia, desta vez, há algo de diferente que merece ser observado: nunca se viram mudanças tão significativas em um espaço tão curto de tempo. Klaus Schwab é defensor da ideia de que a sociedade está vivendo a chamada Quarta Revolução Industrial:¹

As razões por que a nova revolução tecnológica provocará mais agitações do que as revoluções industriais anteriores são aquelas mencionadas na introdução: velocidade (tudo está acontecendo em um ritmo muito mais rápido do que antes), amplitude e profundidade (há muitas mudanças radicais ocorrendo simultaneamente), e a transformação completa de sistemas inteiros.

Outros autores, como Eugênio Facchini Neto e Roberta Scalzilli, apontam que já estamos a caminho da Sociedade 5.0:

Segundo alguns, o termo sociedade 5.0 refere-se à evolução dos modelos anteriores, depois de ter passado de uma sociedade de coletores e caçadores (sociedade 1.0) para uma sociedade de agricultores (sociedade 2.0), evoluindo para uma sociedade industrial (sociedade 3.0) até a nossa sociedade atual da informação (sociedade 4.0). Na sociedade 5.0, a Internet das coisas irá aprofundar ainda mais pessoas e objetos, possibilitando o compartilhamento de todo tipo de conhecimento, mais focada no crescimento sustentável, na integração entre o on-line e o off-line, equilibrando avanço econômico com a

¹ SCHWAB, Klaus. **A quarta revolução industrial**. Tradução Daniel Moreira Miranda. São Paulo: Edipro, 2019. p. 45.

resolução de problemas sociais por meio de um sistema que integra o cyberspaço e o espaço físico.

Nesse contexto, a sociedade 5.0 preconiza o avanço e melhoria da qualidade de vida por meio das inovações tecnológicas, objetivando que humanos e máquinas sejam aliados e não inimigos. Isso ocorre, por exemplo, quando a inteligência artificial, através do uso dos algoritmos, faz recomendações, customizando um produto de acordo com as preferências de cada pessoa. Ou então quando se usam drones para entrega de comidas e medicamentos, carros autônomos, espelho inteligente, robôs assistentes, chatbots, impressoras 3D, entre outras tecnologias que visam alcançar a máxima satisfação e comodidade das pessoas.

Proclama-se que tudo pode vir a ser "inteligente": casas, roupas, automóveis, cidades, indústrias, inclusive a utilização da inteligência artificial em atividades artísticas de criação. Portanto, a ideia de que a criatividade é exclusividade humana está com os dias contados, pois a inteligência artificial já é capaz de compor músicas, peças teatrais e pintar quadros.²

Atualmente, utiliza-se de aplicações para acessar qualquer tipo de informação sobre qualquer assunto que se escolher saber (Google), indicar o melhor trajeto para se chegar a um destino conforme critérios parametrizados e controlados por GPS (Wazze, Google Maps), se relacionar socialmente e profissionalmente (Instagram, Facebook, LinkedIn, TikTok), auxiliar a escolher os melhores filmes e músicas para cada perfil (Netflix, YouTube), solicitar um transporte (Uber), escolher e fazer compras (Amazon, Mercado Livre e outras plataformas), pedir um *delivery* (Ifood), entre outras milhares de funcionalidades (vide *smartphones*, *smartwatches*, etc.).

Destaca-se que, segundo o Fórum Econômico Mundial³, a tecnologia que é a principal responsável por tantos avanços é a inteligência artificial, que atua como um

² FACCHINI NETO, Eugênio; SCALZILLI, Roberta. Pode a ética controlar o desenvolvimento tecnológico? O caso da inteligência artificial, à luz do direito comparado. In: COLOMBO, Cristiano. FALEIROS JR, José Luiz de Moura; ENGELMANN, Wilson. **Tutela jurídica do corpo eletrônico: novos desafios ao direito digital**. Indaiatuba/SP: Foco, 2022. p. 1214. *E-book*.

³ SCHWAB, Klaus. The fourth industrial revolution: what it means, how to respond. **World Economic Forum**, 14 jan. 2016. Disponível em: <https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/>. Acesso em: 12 nov. 2022; GHOSH, Iman. 4 key areas where AI and IoT are being combined. **World Economic Forum**, 15 mar. 2021. Disponível em: <https://www.weforum.org/agenda/2021/03/ai-is-fusing-with-the-internet-of-things-to-create-new-technology-innovations/>. Acesso em: 12 nov. 2022.

"hub" e, quando associada (i) à possibilidade de captação de todo o tipo de informação a respeito do comportamento mundial (*IoT* ou internet das coisas), (ii) à transmissão instantânea desta significativa massa de dados (5G), (iii) os quais podem ser armazenados e classificados com facilidade (*big data*) e (iv) à alta capacidade computacional, permite-se que sejam feitas correlações e previsões com base em estatística gerando uma verdadeira revolução, que constitui a mencionada Quarta Revolução Industrial.

Todavia, Yuval Noah Harari alerta que a inteligência artificial (IA) está entre os maiores desafios da atualidade, exigindo cooperação global para resolver os dilemas oriundos de sua ampla utilização, tendo em vista os diversos riscos sobre a segurança e a ética das decisões tomadas pelas máquinas.⁴ As tecnologias (dentre as quais se insere a inteligência artificial) influenciam em diversas experiências do dia a dia das pessoas e, em virtude disso, alerta-se que o seu uso deve ser refletido, eis que podem afetar garantias dos cidadãos.⁵

Para entender a problemática pesquisada, buscar-se-á compreender sobre a tecnologia envolvida, ou seja, a inteligência artificial.

2.2 Sobre a Inteligência Artificial (IA)

Segundo Eduardo Magrani, Priscilla Silva e Rafael Viola: "A inteligência artificial merece destaque nessa discussão, porque trata justamente da tentativa de se criarem mecanismos capazes de "pensar" e agir de forma cada vez mais autônoma."⁶

Para Marcelo Negri Soares, Marcos Eduardo Kauffman e Kuo-Ming Chao, inteligência artificial é um termo de difícil definição, pois acreditam haver pouco

⁴ HARARI, Yuval Noah. **Homo Deus**: uma breve história do amanhã. São Paulo: Companhia das Letras, 2016.

⁵ DONEDA, Danilo Cesar Maganhoto *et al.* Considerações iniciais sobre inteligência artificial, ética e autonomia pessoal. **Pensar**, Fortaleza, v. 23, n. 4, p. 1-17, out/dez. 2018. DOI: 10.5020/2317-2150.2018.8257. Disponível em: <https://ojs.unifor.br/rpen/article/view/8257>. Acesso em: 05 out. 2022.

⁶ MAGRANI, Eduardo; SILVA, Priscilla; VIOLA, Rafael. Novas perspectivas sobre ética e responsabilidade de inteligência artificial – Parte I – Compreensão da inteligência artificial e dos seus pressupostos de controle e regulação. *In*: FRAZÃO, Ana; MULHOLLAND, Caitlin. *Inteligência artificial e Direito*. São Paulo: Editora Revista dos Tribunais. 2020. posição 2855. *E-book*.

consenso a respeito de seu significado, que muda conforme o recorte temporal, geográfico e cultural dos autores que buscam defini-lo. Por isso, adotam a seguinte interpretação para o termo IA:

Para os fins deste artigo, vamos abordar a inteligência artificial como um ramo da ciência da computação, de certo modo um tanto nebuloso, que busca construir máquinas capazes do que os humanos consideram como um comportamento inteligente. O comportamento inteligente seria o processamento de linguagem similar ao natural, ao humano, com aprendizado de máquina, baseada em algoritmos (cálculos, forma objetiva de resultados) e tomada de decisões (essas, tipicamente subjetivas).⁷

Para Alan Duarte e Ramon de Vasconcelos Negócio, inteligência artificial pode ser definida da seguinte forma:

Com isso, a IA passa a ser um campo de estudo da Ciência da Computação que parte da ideia de que os aspectos da inteligência humana, sobretudo o aprendizado e a resolução de problemas, podem ser precisamente descritos ao ponto de uma máquina ser capaz de simular tais características, tendo como problema central “fazer uma máquina se comportar de maneiras que poderiam ser chamadas de inteligentes se um ser humano assim se comportasse” (McCarthy et al., 1955, p. 11 – tradução livre). Enquanto aplicação prática, a IA pode ser definida como o conjunto de técnicas, métodos e sistemas que objetivam fazer um computador exibir comportamentos inteligentes, sobretudo o aprendizado, a partir da coleta e interpretação de dados externos, a fim de atingir objetivos e tarefas específicas por meio de uma adaptação flexível.⁸

⁷ SOARES, Marcelo Negri; KAUFFMAN, Marcos Eduardo; CHAO, Kuo-Ming. Inteligência artificial: impactos no Direito e na Advocacia. **Revista Direito Público**, Brasília, v. 17, n. 93, 2020. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/3555>. Acesso em: 13 ago. 2022.

⁸ Kaplan e Haenlein (2018) *apud* DUARTE, Alan; NEGÓCIO, Ramon de Vasconcelos. Todos são iguais perante o algoritmo? uma resposta cultural do Direito à discriminação algorítmica. **Revista Direito Público**, Brasília, v. 18, n. 100, 2022. DOI: 10.11117/rdp.v18i100.5869. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/5869>. Acesso em: 4 ago. 2022. p. 227.

Juarez Freitas e Thomas Bellini Freitas interpretam que “inteligência artificial (IA) é um sistema algorítmico adaptável, relativamente autônomo, emulatório da decisão humana”⁹ e a consideram “distinta da operação simbólica incapaz de aprendizagem.”¹⁰ Para os autores, a capacidade de aprendizagem é uma das características nucleares da IA.¹¹

As problematizações para definir o tema inteligência artificial são inúmeras até mesmo para os estudiosos da área da computação. Como será demonstrado a seguir, denota-se que seu conceito também não é unânime nesta área.

Stuart Russel e Peter Norvig¹² explicam o que consideram IA, bem como toda a trajetória desta tecnologia desde a criação do campo de pesquisa e sua evolução, até o momento presente. Os referidos autores mencionam que, historicamente, pesquisadores indicam versões diferentes do que é IA, o que atribuem ao fato de que não há uma convergência sobre o que é considerado inteligência.¹³ Algumas correntes associam o termo inteligência à racionalidade, sendo elas normalmente vinculadas a ciências como matemática, engenharia, estatística e economia; outras correntes associam o conceito de inteligência a comportamento, vinculando-o a áreas como psicologia.¹⁴ Os autores ressaltam que muitos confundem o conceito de IA com o conceito de aprendizagem de máquina, que são distintos.¹⁵

⁹ FREITAS, Juarez; FREITAS, Thomas Bellini. **Direito e Inteligência Artificial**: em defesa do humano. 1. ed. Belo Horizonte: Forum, 2020. p. 30.

¹⁰ *Ibidem*, p. 28.

¹¹ Na visão deles: “O ponto crítico é que os algoritmos de aprendizagem na IA são organizados com maior complexidade do que na automação, pois não apenas seguem regras como também tomam decisões, aprendendo sozinhos sobre os dados coligidos. Implica dizer que, no âmbito da IA, a máquina é um sistema dotado de relativa autonomia. Grife-se: a IA toma decisões. Tal fenômeno não pode ser ignorado”. *Ibidem*, p. 28.

¹² RUSSELL, Stuart.; NORVIG, Peter. **Artificial intelligence**: a modern approach. 4. ed. London: Pearson Education, 2021.

¹³ *Ibidem*, p. 19.

¹⁴ A própria psicologia não tem um conceito único do que se entende por inteligência, sobretudo depois que Howard Gardner desenvolveu a teoria das inteligências múltiplas em 1980 e atualmente defende a existência de pelo menos nove tipos diferentes de inteligência.

¹⁵ “*In the public eye, there is sometimes confusion between the terms “artificial intelligence” and “machine learning.” Machine learning is a subfield of AI that studies the ability to improve performance based on experience. Some AI systems use machine learning methods to achieve competence, but some do not.*” Em tradução livre: Aos olhos do público, às vezes há confusão entre os termos “inteligência artificial” e “aprendizagem de máquina”. O aprendizado de máquina é um subcampo da IA que estuda a capacidade de melhorar o desempenho com base na experiência. Alguns sistemas de IA usam métodos de aprendizado de máquina para obter competência, mas outros não. RUSSELL; NORVIG, *loc. cit.*

Kai Fu Lee explica que, “em meados da década de 1950, os pioneiros da inteligência artificial estabeleceram uma missão com um propósito extremamente ambicioso, mas bem definido: recriar a inteligência humana em uma máquina.”¹⁶

Mas o que se entende por inteligência humana? Assim como Russel e Norvig, Dora Kaufman pontua que “inteligência não é um conceito simples nem consensual, na literatura existem diversas abordagens, nenhuma delas universal.”¹⁷

No Congresso Nacional Brasileiro tramitam alguns projetos de lei com o objetivo de regular a inteligência artificial. Dada a complexidade da discussão, que atualmente ¹⁸ se encontra no Senado Federal, promoveu-se a aglutinação dos projetos de lei existentes e a instalação de uma comissão de juristas, com o fito de promover debates com a sociedade civil e acadêmica para aprimorar as redações elaboradas pelos parlamentares.¹⁹ Após quase 10 meses de intensas discussões e muitos aprendizados, a comissão produziu um relatório e uma nova minuta para a lei, recomendada para votação pelos senadores.²⁰

A futura legislação, se aprovada, definirá inteligência artificial da seguinte maneira:

I – sistema de inteligência artificial: sistema computacional, com graus diferentes de autonomia, desenhado para inferir como atingir um dado conjunto de objetivos, utilizando abordagens baseadas em aprendizagem de máquina e/ou lógica e representação do conhecimento, por meio de dados de entrada provenientes de máquinas ou humanos, com o objetivo de produzir previsões, recomendações ou decisões que possam influenciar o ambiente virtual ou real.²¹

¹⁶ LEE, Kai-Fu. **Inteligência artificial**: como os robôs estão mudando o mundo, a forma como amamos, nos relacionamos, trabalhamos e vivemos. Tradução: Marcelo Barbão. 1. ed. Rio de Janeiro: Globo Livros, 2019. p. 17. *E-book*.

¹⁷ KAUFMAN, Dora. **A inteligência artificial irá suplantar a inteligência humana?** [S.l.]: Estação das Letras e Cores Edi, 2019. posição 136. *E-book*.

¹⁸ Data da atualização: 04/01/2023.

¹⁹ Todo o trabalho produzido pela Comissão de Juristas, incluindo os vídeos de audiências públicas, seminários e contribuições escritas recebidas, podem ser examinado em: CUEVA, Ricardo Villas Bôas *et al.* **Relatório final da Comissão de Juristas responsável por subsidiar elaboração de substitutivo sobre Inteligência Artificial no Brasil**. Brasília: Coordenação de Comissões Especiais, Temporárias e Parlamentares de Inquérito, 2022. Disponível em: <https://legis.senado.leg.br/sdleg-getter/documento?dm=9221643&ts=1671480646036>. Acesso em: 04 jan. 2023.

²⁰ *Ibidem*.

²¹ *Ibidem*, p. 23.

Interpretando os diversos conceitos abordados anteriormente, pode-se verificar que a IA é um grande guarda-chuva de métodos computacionais desenvolvidos para replicar e, em alguns casos, até mesmo a superar ações humanas (tais como raciocínio, fala, reconhecimento, execução de tarefas repetitivas etc.), os quais englobam diversos tipos de abordagem técnicas – cada qual com suas peculiaridades, riscos e seus efeitos colaterais, que merecem ser avaliados e mitigados. Registra-se que é esta a ideia que a pesquisa remete quando usar a terminologia IA ou Inteligência Artificial.

Todavia, dos conceitos expostos anteriormente, faz-se necessário o esclarecimento a respeito de outros dois termos técnicos que foram utilizados para definir IA: algoritmos e aprendizagem de máquina (também conhecida como *machine learning*).

2.3 Algoritmos

Iniciando pelos algoritmos, Ana Frazão os define como “fórmulas ou receitas para execução de tarefas, soluções de problemas, realizações de julgamentos e tomadas de decisões.”²² A autora refere que eles são utilizados para compreender dados e informações do passado na intenção de se tentar prever o futuro.

Duarte e Negócio compreenderam em sua pesquisa que algoritmos se referem a um procedimento de tomada de decisão, descrevendo que “os algoritmos são um sistema de execução de uma programação de código digital, que pode ser determinada por homens e mulheres, assim como por inteligência artificial, isto é, independente de uma programação humana”²³ e atribui relevância a certas informações, conforme preponderância estatística ou programação, o que ocorre de

²² FRAZÃO, Ana. Discriminação algorítmica: compreendendo o que são os julgamentos algorítmicos e o seu alcance na atualidade. Parte I. **JOTA**. 16 jun. 2021. Disponível em: <https://www.jota.info/opiniao-e-analise/colunas/constituicao-empresa-e-mercado/discriminacao-algoritmica-16062021>. Acesso em: 18 nov. 2021.

²³ DUARTE, Alan; NEGÓCIO, Ramon de Vasconcelos. Todos são iguais perante o algoritmo? uma resposta cultural do Direito à discriminação algorítmica. **Revista Direito Público**, Brasília, v. 18, n. 100, 2022. DOI: 10.11117/rdp.v18i100.5869. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/5869>. Acesso em: 4 ago. 2022. p. 225.

forma automatizada a partir da inserção dos dados, do seu tratamento e da emissão dos resultados.

Laura Schertel Mendes e Marcela Mattiuzo também definem o seu conceito de algoritmo.²⁴ Expõem que, apesar de ser frequentemente descrito como sendo um conjunto ou sequência de instruções a serem executadas em uma ordem específica (como uma fórmula), com o intuito de se atingir determinado resultado esperado, as instruções a serem dadas a computadores são diferentes daquelas dadas a seres humanos.

Isso porque os computadores não compreendem conceitos que possuem conotação subjetiva – como exemplo, os termos “suficiente”, “quase” e “ruim”. Para entendê-los, os computadores precisam de métricas objetivas e quantificáveis que, quando identificadas, enquadrarão determinado dado/informação nestas categorias de palavras. Segundo Mendes e Mattiuzo, “é por essa razão que um algoritmo que determine que um celular reduza a luz de sua tela sempre que “quase não haja mais bateria” é inútil. Um computador é capaz de interpretar porcentagens, mas não de determinar o que ‘quase sem bateria’ significa”.

Na sequência, as autoras esclarecem que um dos principais objetivos dos algoritmos é analisar dados, extrair probabilidades e, por meio delas, indicar previsões.²⁵ A crescente disponibilização de dados, captados em maior volume e variedade (*big datas*²⁶), fez com que milhares de decisões fossem automatizadas, impactando a vida de todos.

A definição vai ao encontro do que leciona o pesquisador de ciências da computação Pedro Domingos, que define algoritmo como “uma sequência de instruções que informa ao computador o que ele deve fazer”²⁷, ressaltando que “um

²⁴ MENDES, Laura Schertel; MATTIUZZO, Marcela. Discriminação algorítmica: conceito, fundamento legal e tipologia. **Revista Direito Público**, Brasília, v. 16, n. 90, 2019. Disponível em:

<https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/3766>. Acesso em: 6 ago. 2022.

²⁵ MENDES, Laura Schertel; MATTIUZZO, Marcela. Discriminação algorítmica: conceito, fundamento legal e tipologia. **Revista Direito Público**, Brasília, v. 16, n. 90, 2019. Disponível em:

<https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/3766>. Acesso em: 6 ago. 2022.

²⁶ Quanto às *big datas*, explicam que há três características que podem ter um big data: (i) grande quantidade de dados e de informações coletadas; (ii) possível imprecisão dos dados disponíveis, em razão de seu grande volume e (iii) a busca de correlações em vez de causalidades entre os dados mapeados.

²⁷ DOMINGOS, Pedro. **O algoritmo mestre**: como a busca pelo algoritmo de machine learning definitivo recriará nosso mundo. 1ª edição. São Paulo: Novatec, 2017. p. 20.

algoritmo não é apenas qualquer conjunto de instruções: elas têm de ser suficientemente precisas e não ambíguas para serem executadas por um computador." Como recurso didático, Domingos exemplifica:

uma receita culinária não é um algoritmo porque não especifica exatamente em que ordem as tarefas devem ser executadas ou cada etapa evolvida. Qual seria a medida exata de uma colher de açúcar? Como qualquer pessoa que já tentou preparar uma nova receita sabe, segui-la pode resultar em algo delicioso ou em uma bagunça. Por outro lado, um algoritmo sempre produz o mesmo resultado. Mesmo se uma receita especificar precisamente meio grama de açúcar, isso não resolverá o problema, porque o computador não sabe o que é açúcar ou um grama. Se quiséssemos programar um robô cozinheiro para fazer um bolo, teríamos de instruí-lo a reconhecer o açúcar a partir de um vídeo, a pegar uma colher, e assim por diante. (Ainda estamos trabalhando nisso).²⁸

Assim, pode-se entender que algoritmos são um conjunto de regras pré-definidas que visam atingir determinado(s) objetivo (s). Para sanar as dificuldades enfrentadas durante a sua programação (tais como ensinar os algoritmos a reconhecer o que é açúcar, ou palavras, ou rostos humanos) é que se disseminou o campo da IA que é conhecido como aprendizagem de máquina, que será comentado a seguir.

2.4 Aprendizagem de máquina

Com relação à aprendizagem de máquina, Soares, Kauffman e Chao esclarecem do que se trata:

a aprendizagem de máquina está relacionada com o processo de permitir que computadores aprendam a otimizar algumas tarefas sem o benefício de programação baseada em regras explícitas.²⁹

²⁸ *Ibidem*, p. 21.

²⁹ SOARES, Marcelo Negri; KAUFFMAN, Marcos Eduardo; CHAO, Kuo-Ming. Inteligência artificial: impactos no Direito e na Advocacia. **Revista Direito Público**, Brasília, v. 17, n. 93, 2020. Disponível em:

Facchini e Scalzilli detalham:

Com o passar dos anos, o aperfeiçoamento dos estudos relacionados à inteligência artificial viabilizou o desenvolvimento de algoritmos de *machine learning* (aprendizado de máquina), o que possibilitou que computadores acumulassem conhecimentos para que se reprogramassem automaticamente, por meio de suas próprias experiências. Trata-se, portanto, de uma ciência em que a capacidade de raciocinar logicamente é artificial (não biológica), automatizada e capaz de continuamente aprender, incorporar tal aprendizado e se reprogramar. A técnica de *machine learning* permite que o computador aprenda a evoluir à medida que é exposto a dados ("Big data"), ensejando ações inteligentes baseadas no conhecimento adquirido pelas informações coletadas. Ou seja, é como se a máquina fosse treinada a partir dos dados, "desenvolvendo" a habilidade de aprender e executar uma tarefa", aprendendo inclusive com suas decisões anteriores, como uma criança faz no seu processo de aprendizagem.³⁰

Duarte e Negócio afirmam que a aprendizagem de máquina, também conhecida como *machine learning*, foi uma das principais técnicas responsáveis pela disseminação da inteligência artificial. Para defini-la, os autores se inspiram no conceito de Kevin Murphy e explicam que se trata de "um conjunto de métodos que podem automaticamente determinar padrões em dados, e então usá-los para prever dados futuros ou desempenhar outras formas de tomada de decisão."³¹

Para explicar sobre aprendizagem de máquina (*machine learning*), Domingos esclarece que:

<https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/3555>. Acesso em: 13 ago. 2022.

³⁰ FACCHINI NETO, Eugênio; SCALZILLI, Roberta. Pode a ética controlar o desenvolvimento tecnológico? O caso da inteligência artificial, à luz do direito comparado. In: COLOMBO, Cristiano. FALEIROS JR, José Luiz de Moura; ENGELMANN, Wilson. **Tutela jurídica do corpo eletrônico: novos desafios ao direito digital**. Indaiatuba/SP: Foco, 2022. p. 1214. *E-book*.

³¹ DUARTE, Alan; NEGÓCIO, Ramon de Vasconcelos. Todos são iguais perante o algoritmo? uma resposta cultural do Direito à discriminação algorítmica. **Revista Direito Público**, Brasília, v. 18, n. 100, 2022. DOI: 10.11117/rdp.v18i100.5869. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/5869>. Acesso em: 4 ago. 2022. p. 227.

Todo algoritmo tem uma entrada e uma saída: os dados entram no computador, o algoritmo faz o que precisa com eles, e um resultado é produzido. O *machine learning* faz o contrário: entram os dados e o resultado desejado, e é produzido o algoritmo que transforma um no outro. Os algoritmos de aprendizado – também conhecidos como aprendizes – são aqueles que criam outros algoritmos. Com o *machine learning*, os computadores escrevem seus próprios programas, logo não precisamos mais fazê-lo.³²

A seguir, o autor esclarece que o que alimenta o *machine learning* são os dados utilizados para ensinar o algoritmo e que a popularização da aprendizagem de máquina cresceu nos últimos anos em decorrência da evolução no armazenamento e processamento de dados, os quais podem ser transformados em informações e ações pela IA.³³

Ou seja, os algoritmos programados por *machine learning* aprendem seus trajetos a partir de um *big data* que possui seu grande banco de dados alimentado com informações coletadas a respeito de fatos já ocorridos e registrados no passado.

Harini Suresh e John Guttag fornecem outros detalhes:

*Machine learning is a type of statistical inference that learns, from existing data, a function that can be generalized to new, unseen data. We can find ML algorithms all around us: making personalized Netflix or YouTube recommendations, powering Siri's stilted conversation, providing live transcriptions on our video calls, auto-tagging the people in our photos, deciding whether we are offered job interviews, or approving (or not) tests at the doctor's office. In each of these examples, an ML algorithm has found patterns in a (usually massive) data set and is applying that knowledge to make a prediction about new data points (which might be your photos, medical records, resume, and so on).*³⁴

³² DOMINGOS, Pedro. **O algoritmo mestre**: como a busca pelo algoritmo de machine learning definitivo recriará nosso mundo. 1ª edição. São Paulo: Novatec, 2017. p. 20.

³³ DOMINGOS, Pedro. **O algoritmo mestre**: como a busca pelo algoritmo de *machine learning* definitivo recriará nosso mundo. 1ª edição. São Paulo: Novatec, 2017. p. 26.

³⁴ Em livre tradução: Aprendizagem de máquina é um tipo de inferência estatística que aprende, a partir de dados existentes, uma função que pode ser generalizada para dados novos e não vistos.

Lee comenta que, em seu início, os estudos de IA eram divididos em duas abordagens: a abordagem baseada em regras (também conhecida como sistemas simbólicos ou de especialistas) e a abordagem das redes neurais.³⁵

A primeira abordagem tinha como foco programar computadores a realizar tarefas a partir de uma sequência de regras lógicas, como por exemplo "se X, então Y."³⁶ Contudo, a infinidade de desdobramentos para cada situação limitou a abordagem a jogos com regras definidas, e exigia a consulta a especialistas para orientação de quais decisões deveriam ser tomadas após a ocorrência "X", por exemplo. Neste campo de estudo da inteligência artificial não havia aprendizado de máquina, somente comandos pré-estabelecidos por humanos a serem repetidos pelo computador.

O segundo campo se inspirou no cérebro humano para tentar replicar nosso sistema lógico e, segundo Lee:

ao contrário da abordagem baseada em regras, os construtores de redes neurais em geral não fornecem às redes regras a serem seguidas na tomada de decisões. Eles simplesmente inserem muitos exemplos de um determinado fenômeno – imagens, jogos de xadrez, sons – nas redes neurais e permitem que as próprias redes identifiquem padrões dentro dos dados. Em outras palavras, quanto menos inferência humana, melhor.

As diferenças entre as duas visões podem ser notadas no modo como elas tratam de um problema simples: identificar se existe um gato em uma imagem.

Podemos encontrar algoritmos de ML ao nosso redor: fazer recomendações personalizadas da Netflix ou Youtube, impulsionar a conversa artificial da Siri, fornecer transcrições ao vivo em nossas videochamadas, marcar automaticamente as pessoas em nossas fotos, decidir se nos são oferecidas entrevistas de emprego ou aprovar (ou não) exames no consultório médico. Em cada um desses exemplos, um algoritmo de ML encontrou padrões em um conjunto de dados (geralmente massivo) e está aplicando esse conhecimento para fazer uma previsão sobre novos pontos de dados (que podem ser suas fotos, registros médicos, currículo e assim por diante). *In*: SURESH, Harini; GUTTAG, John. Understanding potential sources of harm throughout the machine learning life cycle. **MIT Case Studies in Social and Ethical Responsibilities of Computing**, 10 ago. 2021. Disponível em: <https://mitserc.pubpub.org/pub/potential-sources-of-harm-throughout-the-machine-learning-life-cycle/release/2?from=7166&to=7866>. Acesso em: 13 nov. 2022.

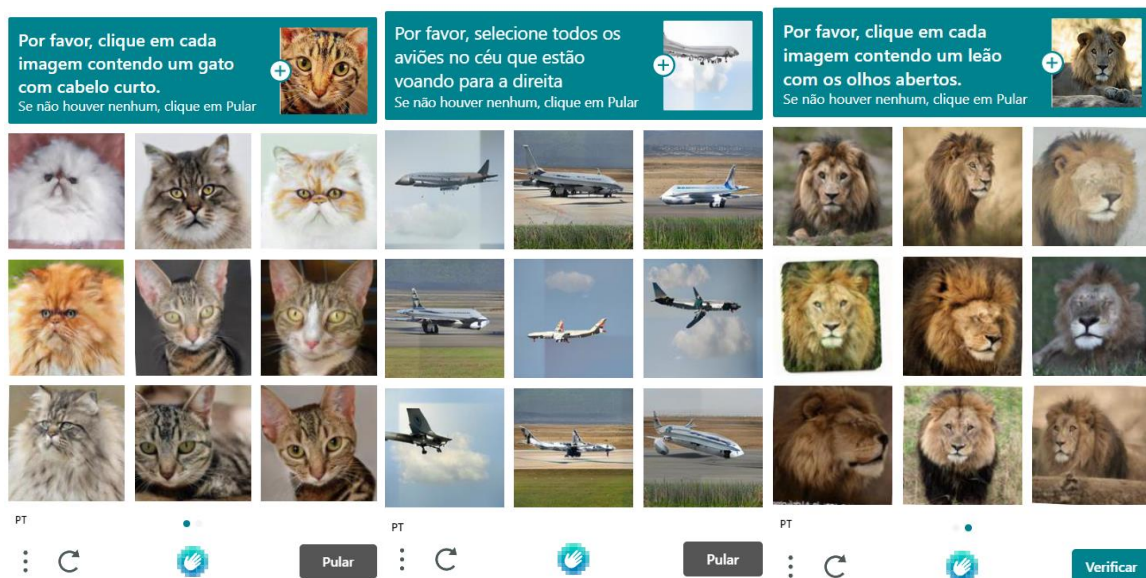
³⁵ LEE, Kai-Fu. **Inteligência artificial**: como os robôs estão mudando o mundo, a forma como amamos, nos relacionamos, trabalhamos e vivemos. Tradução: Marcelo Barbão. 1. ed. Rio de Janeiro: Globo Livros, 2019. p. 17. *E-book*.

³⁶ LEE, Kai-Fu. **Inteligência artificial**: como os robôs estão mudando o mundo, a forma como amamos, nos relacionamos, trabalhamos e vivemos. Tradução: Marcelo Barbão. 1. ed. Rio de Janeiro: Globo Livros, 2019. p. 18. *E-book*.

A abordagem baseada em regras tentaria estabelecer regras nos moldes “se-então” para ajudar o programa a tomar uma decisão: “Se há duas formas triangulares em cima de uma forma circular, então provavelmente há um gato na foto”. A abordagem da rede neural, ao contrário, alimentaria o programa com milhões de amostras de fotos rotuladas como “gato” ou “sem gato”, permitindo que o programa descubra sozinho quais recursos nos milhões de imagens estão mais correlacionados com o rótulo “gato”.³⁷

A título ilustrativo, vale comentar que constantemente vivencia-se a abordagem da rede neural descrita acima na prática, o que ocorre quando um *site* solicita que um humano prove não ser um robô, indicando “identifique as imagens com carros”³⁸ ou ainda quando “ajude a melhorar o Google fotos.”³⁹ Abaixo, na Figura 1, alguns exemplos de rotulação:

Figura 1 – Exemplos de rotulação



Fonte: Compilado feito a partir de imagens do acervo pessoal da autora.

³⁷ *Ibidem*, p. 19.

³⁸ DZIEZA, Josh. Why captchas have gotten so difficult: demonstrating you're not a robot is getting harder and harder. **The Verge**. 01 fev. 2019. Disponível em: <https://www.theverge.com/2019/2/1/18205610/google-captcha-ai-robot-human-difficult-artificial-intelligence>. Acesso em: 15 mar. 2022.

³⁹ VINCENT, James. Google wants you to help train its AI by labeling images in Google Photos: a new optional feature in Google Photos on Android. **The Verge**. 11 nov. 2020. Disponível em: <https://www.theverge.com/2020/11/11/21559930/google-train-ai-photos-image-labelling-app-android-update>. Acesso em: 15 mar. 2022.

Na verdade, além de provar que não há um robô operando o sistema, também se está ensinando a máquina como humanos podem reconhecer estas figuras, ainda que nunca tenham visto aquela imagem específica. A partir das rotulações indicadas por humanos, a máquina “aprende” quais características em comum existem entre as imagens rotuladas como “carro”, por exemplo, e cria sua própria imagem correspondente a este objeto (que também pode ser um caractere, som etc.).

De tal modo, após serem estabelecidas as definições anteriores e esclarecido o conceito de inteligência artificial como uma tecnologia abrangente que engloba diversas abordagens técnicas que devem ser consideradas em eventuais regulações e/ou autorregulações, destaca-se que a pesquisa se concentrou na aprendizagem de máquina, também conhecida como *machine learning (ML)*. Isso porque, atualmente, é a abordagem mais preocupante do ponto de vista de haver dificuldades de controle de seus resultados e, conseqüentemente, de prever seus danos e as ferramentas de mitigação.

Contudo, não se pode esquecer que a aprendizagem de máquina não é sinônimo de inteligência artificial, tendo em vista as diversas outras técnicas capazes de simular aspectos da inteligência humana. Russel e Norvig elencam pelo menos quatro subcampos diferentes de IA⁴⁰, Domingos explica cinco⁴¹, e Nicholas Corrêa⁴² exemplifica pelo menos oito abordagens diferentes de IA que não envolvem aprendizado de máquina. São elas:

We certainly cannot reduce all types of intelligent systems, or “Artificial Intelligence” (AI), to just Machine Learning. We also have the symbolic approach (Newell, 1990), the connectionist approach (Churchland & Sejnowski, 1992), hybrid methodologies (symbolic/connectionist), the mathematical-universal approach (Hutter, 2005), among several other methodologies that seek to develop systems capable of simulating certain cognitive capabilities to solve various

⁴⁰ RUSSELL, Stuart.; NORVIG, Peter. **Artificial intelligence: a modern approach**. 4. ed. London: Pearson Education, 2021. p. 20.

⁴¹ DOMINGOS, Pedro. **O algoritmo mestre: como a busca pelo algoritmo de machine learning definitivo recriará nosso mundo**. 1ª edição. São Paulo: Novatec, 2017.

⁴² CORRÊA, Nicholas Kluge. **Artificial Intelligence ethics and safety: practical tools for creating “good” models**. 2021. p. 4. DOI: 10.13140/RG.2.2.16827.13602. Disponível em: https://www.researchgate.net/publication/357032443_Artificial_Intelligence_Ethics_and_Safety_practical_tools_for_creating_good_models. Acesso em: 13 nov. 2022. p. 4.

*types of problems (e.g., genetic algorithms, dynamic programming, BDI agents, etc.).*⁴³

O avanço da aprendizagem de máquina é um dos motivos que leva a crer que em algumas décadas a inteligência artificial poderá se tornar o que chamam de IA forte, também chamada de IA geral ou "IA nível humano"⁴⁴, que significa que a inteligência artificial poderá atingir o mesmo nível de aprendizagem humana, podendo executar uma tarefa completamente nova, ainda que inicialmente não tenha sido programada especificamente para isso.⁴⁵

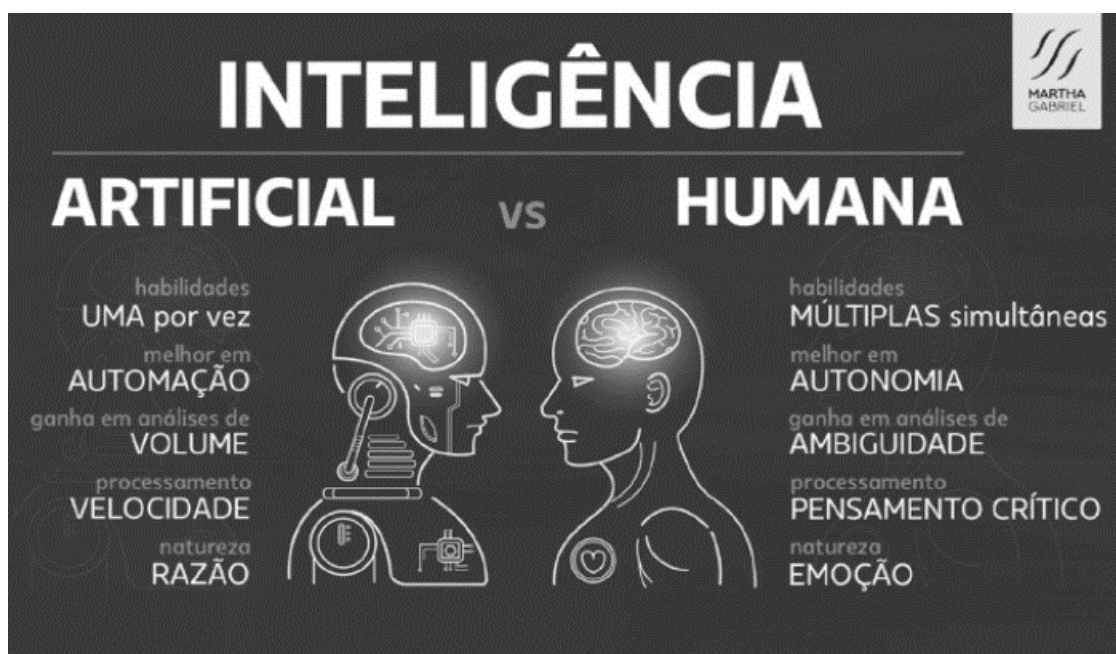
Esta condição ainda não é realidade, eis que as aplicações de inteligência artificial atuam naquilo que foram programadas: por exemplo, um tradutor de textos online traduz textos, ainda que de muitos tipos, mas ele não terá a capacidade de indicar a probabilidade de chuva, caso seja perguntado – a menos que seja programado para isto. Isso é o que se convém chamar de IA fraca, a qual pode até mesmo ultrapassar a inteligência humana em capacidades específicas, mas ainda depende de comandos para atingir novos objetivos. A Figura 2 abaixo demonstra as principais diferenças entre as capacidades da inteligência artificial disponível atualmente e da inteligência humana:

⁴³ Em tradução livre: Certamente que não podemos reduzir todos os tipos de sistemas inteligentes, ou "Inteligência Artificial" (IA), a apenas Aprendizagem de Máquina. Possuímos também a abordagem simbólica (Newell, 1990), o conexionismo (Churchland & Sejnowski, 1992), as metodologias híbridas (simbólica/conexionista), a abordagem matemático-universal (Hutter, 2005), entre diversas outras metodologias que buscam desenvolver sistemas capazes de simular certas capacidades cognitivas para resolver diversos tipos de problemas (e.g., algoritmos genéticos, programação dinâmica, agentes BDI, etc.). *Ibid.*, p. 4.

⁴⁴ Dora Kaufman usa o termo "superinteligência". KAUFMAN, Dora. **A inteligência artificial irá suplantar a inteligência humana?** [S.l.]: Estação das Letras e Cores Edi, 2019. local 826. *E-book*.

⁴⁵ GABRIEL, Martha. **Inteligência artificial: do zero ao metaverso**. 1. ed. Barueri/SP: Atlas, 2022. p. 111.

Figura 2 – Inteligência artificial vs humana



Fonte: Martha Gabriel.⁴⁶

Como se vê, a pretensão é a de tornar a inteligência artificial ainda mais forte, igualando e até ultrapassando a inteligência humana. Domingos explica que quando “pudermos projetar máquinas que sejam mais inteligentes que nós, por sua vez elas devem poder projetar máquinas que também as ultrapassem em inteligência, e isso continuaria *ad infinitum*, deixando a inteligência humana para trás.”, o que é chamado de teoria da singularidade.

Deste modo, diante destas potencialidades que podem contribuir com inúmeros benefícios para o planeta, não se pode olvidar que o uso desmedido e irrestrito da inteligência artificial também pode ocasionar diversos riscos, tais como: concentração de poder em poucas empresas (que passam a deter um monopólio/oligopólio de dados e de tecnologias), invasão de privacidade, injustiças provocadas por discriminação contra grupos minorizados, dificuldade de

⁴⁶ GABRIEL, Martha. **Inteligência artificial**: do zero ao metaverso. 1. ed. Barueri/SP: Atlas, 2022. p. 114.

compreensão acerca do funcionamento da inteligência artificial, dificuldade de responsabilização, entre outros, como será abordado adiante.⁴⁷

2.5 Riscos e aspectos jurídicos da IA

Ao pesquisar sobre a possibilidade de aplicações de IA desenvolverem bases morais autônomas, Paulo Antônio Caliendo Velloso da Silveira reflete que, apesar do estado avançado da tecnologia, até o momento ela corresponde aos comandos humanos e, com isso, reflete os seus dilemas:

A proposta aparentemente simples e compreensível de Asimov não impediria dilemas éticos: e se um robô para proteger um humano, tivesse de causar mal a outro humano? E se o humano causasse mal a si mesmo? Poderia o sistema agir para impedir? Haveria conflito entre o livre-arbítrio humano e o paternalismo artificial? O que significaria "causar dano"? Alimentar-se mal? Qual a extensão desse dever de cuidado ou tutela? As perguntas poderiam suceder em número infinito. Poder-se-ia ainda questionar se seria consistente tratar um robô superinteligente como um mero escravo ou objeto e não como um sujeito de direitos? Digamos que possamos adequadamente listar, para cada pergunta sobre como o algoritmo em IA deveria agir em determinada situação, uma resposta adequada e que cada resposta seja estruturada de modo consistente com as demais respostas. Devemos igualmente pressupor que não somos infalíveis e que não somos oniscientes sobre todos os fatos contingentes. Ou seja, o programa é elaborado por um ser humano para ser aplicado por um sistema artificial inteligente.⁴⁸

⁴⁷ "A nova onda de tecnologias de informação e comunicação (TICs) tornou ainda mais exponencial os possíveis efeitos adversos de uma atividade de tratamento de dados pessoais. Juntas, Internet das Coisas, Big Data e Inteligência Artificial permitem a coleta massiva de informações pessoais e, principalmente, inferências mais intrusivas a respeito dos cidadãos (SPINA, 2014). Com isso, torna-se mais complexo o processo de cognição, avaliação e gerenciamento dos riscos de uma economia de dados. Os agentes de tratamento de dados – controladores e operadores – passaram a deter uma superioridade informacional ainda maior frente aos demais atores – cidadãos e órgãos fiscalizadores – desse ecossistema". *In*: BIONI, Bruno Ricardo; LUCIANO, Maria. O princípio da precaução na regulação de inteligência artificial: seriam as leis de proteção de dados o seu portal de entrada. *In*: MULHOLLAND, Caitlin; FRAZÃO, Ana (Coord.). **Inteligência artificial e direito: ética, regulação e responsabilidade**. São Paulo: Thomson Reuters Brasil, 2019. posição 5697. *E-book*.

⁴⁸ SILVEIRA, Paulo Antônio Caliendo Velloso da. **Ética e Inteligência Artificial: da possibilidade filosófica de agentes morais artificiais**. 1. ed. Porto Alegre: Fi, 2021. v. 1. p. 52.

Uma pesquisa realizada por Danilo Cesar Maganhoto Doneda *et al.*⁴⁹ se prestou a refletir sobre os impactos da IA nos processos de tomada de decisão, concentrando a atenção em como as previsões feitas por IA podem influenciar o acesso de pessoas a bens, serviços e ao mercado.

No estudo, afirmou-se que tais técnicas podem afetar tanto a forma como o indivíduo se percebe no mundo quanto a imagem que a sociedade possa ter dele, o que pode prejudicar sua autonomia, personalidade e moralidade. Destacou-se que, em determinados casos, os algoritmos restringem oportunidades, limitam escolhas econômicas e, na medida em que produzem resultados desiguais, podem violar o princípio da igualdade. Após discorrerem sobre alguns exemplos, como o *credit score* e uso de IA para serviços governamentais, os autores comentam:

Percebe-se, a partir de tais exemplos, o potencial de violação aos direitos fundamentais das decisões automatizadas quando tomadas sem o cumprimento de determinados parâmetros éticos e legais que assegurem a sua transparência e controle individual, a participação do indivíduo no âmbito do processo decisório, bem como a correção e a atualização das informações que servem como input do algoritmo.⁵⁰

Os pesquisadores debatem a autonomia, personalidade jurídica e responsabilidade dos robôs, com destaque para a menção de que estes podem apresentar comportamentos diferentes do que haviam sido previamente programados por seus desenvolvedores, a partir das interações com o ambiente em que estão inseridos.

A seguir, os autores elencam uma série de riscos no uso de IA, sendo eles: redução do controle humano; remoção da responsabilização humana; desvalorização de competências humanas; erosão da autodeterminação humana;

⁴⁹ DONEDA, Danilo Cesar Maganhoto *et al.* Considerações iniciais sobre inteligência artificial, ética e autonomia pessoal. **Pensar – Revista de Ciências Jurídicas**, Fortaleza, v. 23, n. 4, p. 1-17, out/dez. 2018. DOI: 10.5020/2317-2150.2018.8257. Disponível em: <https://ojs.unifor.br/rpen/article/view/8257>. Acesso em: 05 out. 2022.

⁵⁰ DONEDA, Danilo Cesar Maganhoto *et al.* Considerações iniciais sobre inteligência artificial, ética e autonomia pessoal. **Pensar – Revista de Ciências Jurídicas**, Fortaleza, v. 23, n. 4, p. 1-17, out/dez. 2018. DOI: 10.5020/2317-2150.2018.8257. Disponível em: <https://ojs.unifor.br/rpen/article/view/8257>. Acesso em: 05 out. 2022. p. 4.

facilitação de condutas humanas controversas ou mesmo malévolas; preconceito e injustiça.⁵¹

No mesmo sentido, Ana Frazão diagnostica aspectos prejudiciais no uso de IA⁵², relatando que o atual contexto econômico é movido a dados e o capitalismo é de vigilância, em que são utilizados algoritmos para análises complexas, tais como decisões e diagnósticos que para serem eficazes são altamente invasivos, provocando uma devassa na intimidade das pessoas e impactando nas possibilidades e acessos a direitos e oportunidades. Exemplos não faltam: quando decidem quem conseguirá obter crédito, qual taxa de juros irá pagar, quem será recrutado para determinada vaga de emprego, a probabilidade de reincidência de réus, quem será atropelado por carros autônomos em caso de acidente inevitável, dentre outras situações. Segundo a autora:

E a consequência disso é uma perda – não um ganho – de liberdade, já que tais práticas procuram moldar e predizer o comportamento dos indivíduos de acordo com trajetórias de oportunidades e desejos que são determinadas externamente. Consequentemente, os algoritmos preocupam tanto quando acertam como quando erram. Preocupam quando acertam, pois podem revelar aspectos íntimos da personalidade das pessoas que elas gostariam de manter em segredo, até para não serem tolhidas em seus direitos e oportunidades em razão disso. Os algoritmos preocupam também quando erram, pois desconfiguram a individualidade, atribuindo às pessoas características que elas não têm e que também podem ser utilizadas para lhes tolher direitos e oportunidades, com o agravante de que tais decisões são baseadas em juízos totalmente equivocados.⁵³

⁵¹ Na mesma pesquisa, os autores também elencam benefícios propiciados pelo uso de IA, tais como: cuidados de saúde; acessibilidade; agricultura e meio ambiente; transporte e outros. Todavia, este tópico não está em discussão na presente pesquisa.

⁵² FRAZÃO, Ana. Responsabilidade civil dos administradores de sociedades empresárias por decisões tomadas com base em sistemas de inteligência artificial. In: FRAZÃO, Ana; MULHOLLAND, Caitlin (Coord.). **Inteligência artificial e Direito: ética, regulação e responsabilidade**. 2. ed. rev. atual. e ampl. São Paulo: Thomson Reuters Brasil, 2020. p. 484. *E-book*.

⁵³ FRAZÃO, Ana. Responsabilidade civil dos administradores de sociedades empresárias por decisões tomadas com base em sistemas de inteligência artificial. In: FRAZÃO, Ana; MULHOLLAND, Caitlin (Coord.). **Inteligência artificial e Direito: ética, regulação e responsabilidade**. 2. ed. rev. atual. e ampl. São Paulo: Thomson Reuters Brasil, 2020. p. 487. *E-book*.

Atentos aos potenciais danos que podem ser causados, Fabrício Bon Vecchio e Débora Manke Vieira investigam se a constante delegação de decisões empresariais a sistemas de inteligência artificial (cujo processo algorítmico nem sempre é claro e é passível de causar danos) pode caracterizar negligência, bem como quem deverá responder por eventuais prejuízos ocasionados. Destacam que se por um lado a responsabilização é necessária, de outro ela pode impactar no crescimento econômico ao aumentar os riscos da atividade empresarial.⁵⁴

Os autores também identificam controvérsias no uso de IA, destacando casos de empresas que disponibilizaram algoritmos racistas e preconceituosos para seleção de currículos, contratos inteligentes com parâmetros equivocados etc. Sublinham que uma decisão errônea pode ocasionar uma punição ou danos à reputação, passíveis de originar até mesmo a falência da empresa. Referem que, mesmo que cuidados sejam tomados, é preciso se ter em mente que será pouco provável o impedimento de todos os riscos.⁵⁵

Miriam Wimmer também reforça sobre os crescentes questionamentos que surgem a partir do uso de IA para tomadas de decisões, decorrentes da “percepção de que tais sistemas poderão não apenas ser imprevisíveis quanto aos seus efeitos gerais sobre a sociedade, como também poderão agir fora do escopo da responsabilidade humana.”⁵⁶ A autora destaca que a crescente autonomia e complexidade da IA coloca em xeque o controle humano, ao passo que nem sempre se compreende, se prevê ou se controla o resultado dos algoritmos.

Wimmer elenca questões jurídicas como “reforço de estereótipos, ao uso de dados pessoais para formação de perfis comportamentais, à ausência de sensibilidade para contexto e cultura, dentre outras.”⁵⁷ e afirma que, no que toca a

⁵⁴ VECCHIO, Fabrício Bon; VIEIRA, Débora Manke. O controle societário, a teoria Business Judgment Rule e a influência da Inteligência Artificial. In: VEIGA, Fábio da Silva; ZALUCKI, Mariusz (Org.). **Legaltech, Artificial Intelligence and the future of legal practice**. Porto/Kraków: Instituto Iberoamericano de Estudos Jurídicos; AFM Kraków University, 2022. p. 293-301.

⁵⁵ VECCHIO; VIEIRA, *loc. cit.*

⁵⁶ WIMMER, Miriam. Responsabilidade de agentes empresariais por ilícitos administrativos praticados por sistemas de inteligência artificial. In: FRAZÃO, Ana; MULHOLLAND, Caitlin (Coord.). **Inteligência artificial e direito: ética, regulação e responsabilidade**. 2. ed. ver. atual. e ampl. São Paulo: Thomsom Reuters Brasil, 2020. p. 10135. *E-book*.

⁵⁷ WIMMER, Miriam. Responsabilidade de agentes empresariais por ilícitos administrativos praticados por sistemas de inteligência artificial. In: FRAZÃO, Ana; MULHOLLAND, Caitlin (Coord.). **Inteligência artificial e direito: ética, regulação e responsabilidade**. 2. ed. ver. atual. e ampl. São Paulo: Thomsom Reuters Brasil, 2020. p. 10274. *E-book*.

responsabilidade jurídica, deve-se problematizar a opacidade dos processos decisórios, a qual dificulta a análise de responsabilidade subjetiva, tais como a verificação da existência de voluntariedade ou culpabilidade.

Eduardo Magrani, Priscilla Silva e Rafael Viola mencionam que "agentes são capazes de influenciar as relações entre as pessoas, moldando comportamentos e visões de mundo, especialmente quando parte do seu funcionamento goza de alta complexidade tecnológica e autonomia"⁵⁸, e acompanham a conclusão de que não há respostas seguras sobre como lidar com os potenciais prejuízos oriundos de erros de programação ou resultantes das técnicas de aprendizagem de máquina que podem incorporar condutas indesejadas não previstas por seus programadores. Nesse sentido, prescrevem que tão importante quanto desenvolver a tecnologia é estabelecer fundamentos éticos mínimos para que ela opere.⁵⁹

Bruno Ricardo Bioni e Maria Luciano atentam que algoritmos nem sempre são isentos de subjetividades, erros ou manipulação, eis que os dados importados são selecionados por alguém, assim como o que fazer com eles é prescrito previamente. Relatam a frequência com que tem havido problemas com algoritmos enviesados, o que demonstra o abismo existente entre os desenvolvedores de IA e aqueles que serão impactados por esta tecnologia.⁶⁰

Como se relatou acima, pesquisadores apontam os inúmeros riscos⁶¹ do uso de inteligência artificial na automatização de decisões, mesmo nas mais simples do

⁵⁸ MAGRANI, Eduardo; SILVA, Priscilla; VIOLA, Rafael. Novas perspectivas sobre ética e responsabilidade de inteligência artificial – Parte I – Compreensão da inteligência artificial e dos seus pressupostos de controle e regulação. In: FRAZÃO, Ana; MULHOLLAND, Caitlin. **Inteligência artificial e Direito**. São Paulo: Revista dos Tribunais. 2020. posição 3230. *E-book*.

⁵⁹ *Ibidem*, posição 2988. *E-book*.

⁶⁰ BIONI, Bruno Ricardo; LUCIANO, Maria. O princípio da precaução na regulação de inteligência artificial: seriam as leis de proteção de dados o seu portal de entrada. In: MULHOLLAND, Caitlin; FRAZÃO, Ana (Coord.). **Inteligência artificial e direito: ética, regulação e responsabilidade**. São Paulo: Thomson Reuters Brasil, 2019. *E-book*.

⁶¹ Para Doneda *et al.*: a autopercepção dos indivíduos no mundo, a imagem que a sociedade pode ter de cada um, a autonomia, a personalidade, a moralidade, o acesso das pessoas a bens, a serviços e aos mercados, restrição de oportunidades, limitação de escolhas econômicas, produção de resultados desiguais, desigualdade, violação de direitos fundamentais, ausência de transparência e de controle individual, exclusão do indivíduo do processo decisório, impossibilidade e correção das informações que alimentam algoritmo, imprevisibilidade dos resultados com relação ao que havia sido programado, privacidade, confiança, discriminação, responsabilidade, redução do controle humano, remoção da responsabilidade humana, desvalorização das competências humanas, erosão da autodeterminação humana, facilitação de condutas humanas controversas, preconceito e injustiça.

dia a dia, em razão dos impactos econômicos e sociais que podem ocasionar na vida de diversas pessoas.

De todos os potenciais danosos verificados, um dos mais citados é o risco desafiador de ocasionar dificuldades para exercício de direitos por certos grupos de indivíduos, o que se convencionou chamar de discriminação algorítmica.

O mencionado risco se trata de um efeito colateral cuja origem, em alguns casos, advém dos dados e informações disponíveis a respeito da sociedade atual, refletindo vieses que são históricos. Nestes casos, ainda que a programação esteja correta e que os dados sejam fidedignos, as decisões automatizadas podem replicar em larga escala os preconceitos e a desigualdade cultural do contexto social que deu origem aos dados.

Soma-se a isto o fato de que, no estado da arte da tecnologia, nem sempre é possível compreender o raciocínio adotado pelo algoritmo, tendo em vista que por vezes seus trajetos e correlações foram tantos e tão complexos, que ultrapassam a atual capacidade de compreensão humana, mesmo quando esmiuçado por técnicos.⁶²

Para Ana Frazão: invasão de privacidade, devassa na intimidade, impacto nas possibilidades e acessos a direitos e oportunidades, perda de liberdade, moldagem e predição do comportamento de indivíduos de acordo com trajetórias de oportunidades e desejos que são determinadas externamente, revelação de aspectos íntimos da personalidade das pessoas que gostariam de manter em segredo, desconfiguração da individualidade, atribuição às pessoas de características que elas não têm, juízos equivocados.

Para Vecchio e Vieira: falta de clareza, danos, racismo, preconceito, parâmetros equivocados, possibilidades de punição e danos à reputação (para empresas).

Para Miriam Wimmer: imprevisibilidade quanto aos efeitos gerais sobre a sociedade e quanto ao resultado dos algoritmos, ação fora do escopo da responsabilidade humana, risco ao controle humano, reforço a estereótipos, uso de dados pessoais para formação de perfis comportamentais, ausência de sensibilidade para contexto e cultura, opacidade dos processos decisórios.

Para Magrani, Silva e Viola: influência as relações entre as pessoas, molda comportamentos e visões de mundo, ausência de respostas seguras, erros de programação ou resultantes das técnicas de aprendizagem de máquina que podem incorporar condutas indesejadas não previstas por seus programadores.

Para Bioni e Luciano: nem sempre são isentos de subjetividades, erros ou manipulação, algoritmos enviesados, abismo existente entre os desenvolvedores de IA e aqueles que serão impactados por esta tecnologia.

⁶² Esta dificuldade de compreensão é constantemente chamada pela doutrina de 'opacidade'. Também é comum referirem que o algoritmo se trata de uma *black box* (ou caixa preta). Sobre o ponto, a pesquisa de Miriam Wimmer sistematiza: Nesse sentido, Burrell (2016, p. 2) identifica três categorias de opacidade algorítmica: (i) opacidade intencional, como mecanismo corporativo ou institucional de autoproteção e ocultação; (ii) opacidade decorrente do fato de que escrever e ler código computacional é uma habilidade limitada a especialistas; e (iii) opacidade que resulta do descasamento entre os procedimentos matemáticos de algoritmos capazes de aprendizado e os estilos humanos de interpretação semântica, o que torna os algoritmos opacos até mesmo para os

Além disso, ainda que seja possível compreender a lógica algorítmica que levou a determinado resultado, nem sempre as empresas desenvolvedoras permitem uma auditoria em seus sistemas (ainda mais quando esta auditoria possa gerar provas em seu desfavor), e em determinados casos elas estão protegidas por regras de segredo industrial.

Mas não é só: diferente do que ocorre com os preconceitos humanos, que podem ser diversos (sociais/culturais) e manifestados de forma pulverizada, uma decisão tomada por IA pode atingir uma escala muito grande de pessoas, pois o volume que ela opera é exponencial. Sendo assim, em vez de termos uma pluralização de tendências e tensões que se contrabalanceiam, teremos apenas um modelo padrão preconceituoso aplicado a todos em grande escala, refletindo a escolha de quem o criou.

Em virtude destas complexidades, que por vezes são contrárias à intenção dos próprios desenvolvedores e usuários da tecnologia, é que a pesquisa passará a abordar especificamente a discriminação algorítmica, para, na sequência, investigar se há instrumentos legais e ferramentas práticas que podem/devem ser utilizadas para sua mitigação.

2.6 Discriminação algorítmica

Na tentativa de solucionar afrontas ao princípio da igualdade, para além das medidas constitucionais e/ou legais, pode-se acreditar que algoritmos são importantes aliados na redução de decisões discriminatórias, na medida que agirão por meio de condições pré-estabelecidas na análise de características determinadas, sem olhar a quem se dirigem as decisões tomadas, promovendo um tratamento mais igualitário.

Todavia, como visto anteriormente, há inúmeros casos em que ocorre exatamente o contrário. Em vez de promover igualdade, algoritmos de IA que não forem programados com o devido cuidado e diligência podem promover ainda mais

seus programadores." WIMMER, Miriam. Responsabilidade de agentes empresariais por ilícitos administrativos praticados por sistemas de inteligência artificial. In: FRAZÃO, Ana; MULHOLLAND, Caitlin (Coord.). **Inteligência artificial e direito: ética, regulação e responsabilidade**. 2. ed. ver. atual. e ampl. São Paulo: Thomsom Reuters Brasil, 2020. p. 10706. *E-book*.

desigualdade e reforçar preconceitos – fatos que precisam ser mitigados assim como se estivessem sendo cometidos por qualquer pessoa humana.

Como será descrito adiante, há exemplos de decisões discriminatórias tomadas por aplicações de inteligência artificial decorrentes da reprodução de vieses históricos presentes na humanidade, tais como preconceitos e dominação social conforme o gênero, a etnia ou a ancestralidade, a cultura, a religião etc. Contudo, tais segregações representam clara violação a direitos humanos, positivados em nossa Constituição Federal como direitos fundamentais.

Laura Schertel Mendes e Marcela Mattiuzzo introduzem o tema da discriminação algorítmica contextualizando que a evolução tecnológica provocou importantes contribuições para os processos de tomada de decisões, auxiliando a redução de riscos. Contudo, exemplificam:

Nos negócios de credit report e credit score, a automatização dos processos decisórios foi, em um primeiro momento, vista como um meio de superar as tendências de enviesamento e discriminação. No entanto, logo tornou-se claro que o método estatístico, o qual teoricamente receberia dados objetivos como inputs e, portanto, deveria gerar resultados objetivos como outputs, poderia reproduzir vieses já existentes, levando também a resultados discriminatórios. Isso porque, em primeiro lugar, nexos de causalidade e correlações são, muitas vezes, predefinidos pelos controladores dos dados, que, por sua vez, transmitem aos algoritmos os mesmos vieses presentes nos processos tradicionais de tomada de decisões.

Ou seja, se alguém acredita que as mulheres são inapropriadas de modo geral para alguns tipos de atividade – por exemplo, para a engenharia mecânica – e essa pessoa programa um algoritmo que internaliza tal lógica, o output de tal algoritmo poderá apresentar essas mesmas inclinações, independentemente da qualidade do input. Mesmo em casos em que o algoritmo seja programado para identificar suas próprias correlações a partir da colheita de dados brutos já existentes – o que deveria eliminar o problema de transferência de predisposições do programador –, ainda assim poderia acabar reproduzindo correlações discriminatórias presentes em tais dados. Em outras palavras, os algoritmos poderiam absorver padrões discriminatórios presentes na sociedade

e replicá-los como uma “verdade objetiva”. Ou seja, mesmo que o designer do algoritmo não acredite que homens seriam engenheiros mecânicos melhores que mulheres, em havendo no conjunto de dados analisado elementos suficientes a indicar que o gênero pode ser uma variável relevante para determinar tais aptidões – por conta do maior número de homens do que mulheres no ramo da engenharia, por exemplo –, o output poderia reproduzir as condições discriminatórias existentes ao invés de auxiliá-las.⁶³

Ao investigar o tema, Mendes e Mattiuzzo ilustram que discriminação costuma remeter à ideia de exclusão de uma pessoa ou um grupo pelo fato de possuir alguma(s) característica(s) e relatam que para fins de sua pesquisa, a discriminação objeto de foco foi:

os resultados discriminatórios que decorrem do fato de alguém pertencer a determinado grupo e ser julgado a partir das características desse grupo; um cenário no qual as características individuais de uma pessoa são desconsideradas, e aquela pessoa é vista somente como um membro de um dado conjunto de pessoas.⁶⁴

Aprofundando a investigação, as pesquisadoras explicam sobre os dois principais tipos de generalizações, as consistentes e as inconsistentes.

As consistentes podem ser universais, não universais ou comparativas. A universal ocorre quando uma afirmação se aplica a 100% das pessoas (exemplo utilizado: humanos são mortais). A não universal se enquadra quando uma afirmação descreve a maioria de determinado grupo; e, a comparativa é utilizada quando uma determinada informação/característica é mais presente em um grupo do que em outro, mas não necessariamente reflete universalmente ou

⁶³ Barocas e Selbst (2016) *apud* MENDES, Laura Schertel; MATTIUZZO, Marcela. Discriminação algorítmica: conceito, fundamento legal e tipologia. **Revista Direito Público**, Brasília, v. 16, n. 90, 2019. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/3766>. Acesso em: 6 ago. 2022. p. 40; 41.

⁶⁴ MENDES, Laura Schertel; MATTIUZZO, Marcela. Discriminação algorítmica: conceito, fundamento legal e tipologia. **Revista Direito Público**, Brasília, v. 16, n. 90, 2019. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/3766>. Acesso em: 6 ago. 2022. p. 47.

majoritariamente a amostra de nenhum dos grupos. Por generalizações inconsistentes, tem-se afirmações que não são universais ou não se aplicam à maioria e nem são predominantes quando comparados com outros grupos. Para ilustrar este último exemplo, as autoras relembram a tese de Cesare Lombroso sobre as características físicas do "homem delinquente". Em resumo (Quadro 1):

Quadro 1 – Tipos de generalizações

Consistentes		Inconsistentes
Universais	Quando uma afirmação se aplica a 100% das pessoas (exemplo utilizado: humanos são mortais).	Afirmações que não são universais ou não se aplicam à maioria e nem são predominantes quando comparados com outros grupos. Para ilustrar, relembram a tese de Cesare Lombroso sobre as características físicas do "homem delinquente".
Não universais	Quando uma afirmação descreve a maioria de determinado grupo.	
Comparativas	Quando uma determinada informação/característica é mais presente em um grupo do que em outro, mas não necessariamente reflete universalmente ou majoritariamente a amostra de nenhum dos grupos.	

Fonte: Elaborado pela autora com base Mendes e Mattiuzzo.⁶⁵

Assim, as autoras esclarecem que apesar de que as generalizações possam aparentar ser negativas, por desconsiderar as individualidades de cada um, as discriminações para fins atuariais são amplamente utilizadas no mundo jurídico e auxiliam em muitas tomadas de decisão – e para explicar, as autoras mencionam o critério definido para maioria (que é o mesmo para todos, independentemente da maturidade individual de cada um) e para regras de trânsito (cuja velocidade máxima é sempre um limite objetivo, independentemente das capacidades de cada motorista

⁶⁵ MENDES, Laura Schertel; MATTIUZZO, Marcela. Discriminação algorítmica: conceito, fundamento legal e tipologia. *Revista Direito Público*, Brasília, v. 16, n. 90, 2019. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/3766>. Acesso em: 6 ago. 2022.

e de cada veículo). Demonstram assim que é comum a utilização das características de um grupo para avaliar a todos e que tais decisões são tomadas por meio de *proxies*⁶⁶ (a idade biológica pode ter correlação com a maturidade mental, porém, se cada indivíduo fosse avaliado, o atingimento da maioridade poderia ocorrer em idades diferentes).

Neste sentido, esclarecem que as discriminações estatísticas podem ser endógenas (vinculada a uma variável interna que pode integrar a mesma diferenciação. Exemplo: mulheres foram historicamente afastadas das oportunidades profissionais em razão de serem responsáveis pelas tarefas domésticas. Tais dados/informações podem levar a menos oportunidades para mulheres); ou exógenas (quando a variável é externa, exemplo: seguros automotivos costumam ser mais caros para homens jovens, pois possuem histórico de maior envolvimento em acidentes. O fato de pagarem mais caro não fará com que se envolvam em mais acidentes). O Quadro 2 a seguir sistematiza estas classificações:

Quadro 2 – Discriminações algorítmicas estatísticas

	Endógenas	Exógenas
Descrição	Quando a variável que distingue um grupo do outro é interna e pode até mesmo integrar a diferenciação.	Quando a variável que distingue um grupo do outro é externa.
Exemplo	Mulheres foram historicamente afastadas das oportunidades profissionais em razão de serem responsáveis pelas tarefas domésticas. Tais dados/informações em menor quantidade nos sistemas, podem levar a menos oportunidades para mulheres.	Seguros automotivos costumam ser mais caros para homens jovens, pois possuem histórico de maior envolvimento em acidentes. O fato de pagarem mais caro não fará com que se envolvam em mais acidentes.

Fonte: Elaborado pela autora com base em Mendes e Mattiuzzo.⁶⁷

⁶⁶ O conceito de *proxies* (plural de *proxy*) será detalhado adiante, mas pode ser entendido como uma variável que, a princípio, não tem vínculo com o problema, mas que acaba por ser utilizada como uma correlação indicativa da informação que se precisaria obter.

⁶⁷ MENDES, Laura Schertel; MATTIUZZO, Marcela. Discriminação algorítmica: conceito, fundamento legal e tipologia. **Revista Direito Público**, Brasília, v. 16, n. 90, 2019. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/3766>. Acesso em: 6 ago. 2022.

Assim, demonstram que a discriminação algorítmica endógena pode ser ineficiente e promover o efeito *feedback loop*, que é a repetição de padrões encontrados nos dados – tais padrões refletem resultados que gerarão novos dados que retroalimentarão o sistema, reforçando uma situação.

Ato seguinte, as pesquisadoras conceituam o que entendem por discriminação algorítmica:

o termo “discriminação algorítmica” é utilizado, neste artigo, para englobar tanto cenários que envolvem afirmações estatisticamente inconsistentes quanto cenários em que as afirmações, embora estatisticamente lógicas, de alguma forma tomam os indivíduos que dela são objeto não de forma efetivamente individualizada, mas apenas como parte de um grupo. Isso porque, a nosso ver, uma classificação, ainda que consistente sob o ponto de vista estatístico, pode em alguns casos se mostrar injusta.⁶⁸

Duarte e Negócio também se propuseram a analisar a discriminação algorítmica e entender como o direito deve buscar soluções para evitar tais ofensas a grupos de indivíduos.

Os autores partem de fatos para elucidar que nem sempre os algoritmos serão neutros, pois refletem em sua programação vieses, crenças e equívocos humanos, eis que desenvolvidos por seres humanos falíveis. Mencionam que os próprios dados que são utilizados para treinar os algoritmos podem ser enviesados ou tendenciosos, na medida em que refletem aspectos da sociedade que são assim.

Compreendem os autores que este novo tipo de discriminação soa mais sofisticado, pois ausente a intencionalidade dos agentes, merecendo uma atualização das bases jurídicas para contemplá-lo. Assim, adentram ao fundamento da não discriminação, que é o princípio da igualdade. Discorrem sobre a evolução deste direito, que em sua primeira geração era visto como uma igualdade formal, que tornava todos iguais perante a lei, independentemente de classe social, evoluindo após para uma igualdade material e fundamental, que embasou e foi complementada pelo direito à diversidade, à diferença.

⁶⁸ *Ibidem*, p. 51.

Os pesquisadores referem que o conceito de discriminação indica uma ofensa ao princípio da isonomia, em virtude dos prejuízos sociais gerados para os destinatários dos tratamentos desiguais. Ambos ressaltam que a discriminação, por si, não é ruim ou nociva e em alguns casos pode ser até mesmo necessária para efetivar o próprio princípio da igualdade, esclarecendo que o tratamento desigual pode ocorrer quando justificado por princípios contrapostos e que “busca-se, com as normas antidiscriminatórias, a eliminação de discriminações que mitiguem, eliminem ou prejudiquem, em alguma medida, o exercício de direitos fundamentais de modo injusto.”⁶⁹

Neste sentido, argumentam que o direito deve proteger os indivíduos para não serem impedidos de se desenvolverem socialmente, registrando que o “princípio da igualdade, por conseguinte, só pode ser um princípio jurídico eficaz na medida em que as instituições sociais identificam os processos responsáveis pelas diferentes formas de discriminação.”⁷⁰ Os autores também comentam que, embora haja diversas normas no sentido de obstar a discriminação negativa de indivíduos, sua ocorrência é institucional e estrutural, não acontecendo sempre de maneira direta (com intencionalidade e arbitrariedade):

A ideia de discriminação engloba como prática qualquer distinção, exclusão, restrição ou preferência, cuja finalidade seja impedir que os indivíduos desfrutem, em pé de igualdade, de direitos fundamentais, ainda que referidas [discriminações] ocorram despidas de intencionalidade. Ou seja, a discriminação está não só na causalidade, mas também nas correlações.⁷¹

⁶⁹ DUARTE, Alan; NEGÓCIO, Ramon de Vasconcelos. Todos são iguais perante o algoritmo? uma resposta cultural do Direito à discriminação algorítmica. **Revista Direito Público**, Brasília, v. 18, n. 100, 2022. DOI: 10.11117/rdp.v18i100.5869. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/5869>. Acesso em: 4 ago. 2022. p. 227.

⁷⁰ *Ibidem*.

⁷¹ Tischbirek (2019) *apud* DUARTE, Alan; NEGÓCIO, Ramon de Vasconcelos. Todos são iguais perante o algoritmo? uma resposta cultural do Direito à discriminação algorítmica. **Revista Direito Público**, Brasília, v. 18, n. 100, 2022. DOI: 10.11117/rdp.v18i100.5869. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/5869>. Acesso em: 4 ago. 2022. p. 224.

Dando sequência ao tema, os autores atentam:

A esfera pública – com a mudança dos paradigmas dos meios de comunicação, isto é, tornando a internet o principal meio de comunicação da sociedade – passa a ser mediada não somente pela imprensa tradicional como também por atores privados que administram intermediários e aplicativos da internet, o que os vincula a um compromisso de inclusão pela administração de sua normatividade técnica. Assim, a garantia da igualdade pela heterogeneidade é um compromisso mais amplo que inclui e vincula esses agentes privados e sua programação de algoritmos.⁷²

Conforme mencionado por Duarte e Negócio, o princípio da igualdade entre os cidadãos é uma das primeiras premissas do Estado Brasileiro. O princípio da igualdade, além de estar expressamente positivado na Constituição Federal Brasileira de 1988 e outros diplomas legais, também foi examinado em suas dimensões por Ingo Wolfgang Sarlet e Gabrielle Sarlet:

Igualdade em sentido material, além disso, significa proibição de tratamento arbitrário, ou seja, a vedação da utilização, para o efeito de estabelecer as relações de igualdade e de desigualdade, de critérios intrinsecamente injustos e violadores da dignidade da pessoa humana, de tal sorte que a igualdade, já agora na segunda fase de sua compreensão na seara jurídico-constitucional, opera como exigência de critérios razoáveis e justos para determinados tratamentos desiguais.

[...]

Com efeito, na CF, objeto imediato de nossa atenção, a igualdade obteve lugar de acentuado destaque em várias passagens do texto constitucional, [...]. Além disso, a igualdade se apresenta no texto constitucional tanto como princípio estruturante do próprio Estado Democrático de Direito quanto na condição de norma impositiva de tarefas para o Estado, bastando, neste contexto, referir o disposto no art. 3º, que, no âmbito dos objetivos fundamentais (com destaque para os incisos III e IV), elenca a redução das desigualdades regionais e a

⁷² *Ibidem*, p. 225.

promoção do bem de todos, sem preconceitos de origem, de raça, de sexo, de cor, de idade e de quaisquer outras formas de discriminação.

Assim como se deu em outras ordens constitucionais contemporâneas, também a CF não se limitou a enunciar um direito geral de igualdade, como ocorreu no art. 5º, caput [...], mas estabelece, ao longo do texto, uma série de disposições impositivas de um tratamento igualitário e proibitivas de discriminação [...].

Nessa perspectiva, é possível afirmar que também no Brasil o princípio (e direito) da igualdade abrange pelo menos três dimensões: a) proibição do arbítrio, de modo que tanto se encontram vedadas diferenciações destituídas de justificação razoável com base na pauta de valores constitucional, quanto proibido tratamento igual para situações manifestamente desiguais; b) proibição de discriminação, portanto, de diferenciações que tenham por base categorias meramente subjetivas; c) obrigação de tratamento diferenciado com vistas à compensação de uma desigualdade de oportunidades, o que pressupõe a eliminação, pelo Poder Público, de desigualdades de natureza social, econômica e cultural.⁷³

Ingo Wolfgang Sarlet ainda ressalva que o princípio da igualdade deve ser respeitado não somente nas ações entre o Estado com seus cidadãos, mas também nas relações entre particulares, especialmente nas proibições de discriminação:

Também o direito geral de igualdade e o correspondente princípio da isonomia (igualdade) encontram-se diretamente ancorados na dignidade da pessoa humana, não sendo por outro motivo que a Declaração Universal da ONU consagrou que todos os seres humanos são iguais em dignidade e direitos. Assim, constitui pressuposto essencial para o respeito da dignidade da pessoa humana a garantia da isonomia entre todos os seres humanos, que, portanto, não podem ser submetidos a tratamento discriminatório e arbitrário, razão pela qual intoleráveis a escravidão, a discriminação racial, perseguições por motivos de religião, gênero, orientação sexual, enfim, toda e qualquer ofensa ao princípio isonômico na sua dupla dimensão formal e material.

⁷³ SARLET, Ingo Wolfgang; SARLET, Gabrielle. Igualdade como proibição de discriminação e direito à (e dever de) inclusão: o acesso ao ensino superior e a regulamentação do estatuto brasileiro das pessoas com deficiência. **Revista Direito Público**, Brasília, v. 14, n. 78, 2017, 197-226 nov./dez. 2017. p. 204-205.

[...]

A compreensão do princípio da igualdade e correspondentes direitos de igualdade na perspectiva (material) da dignidade da pessoa humana assume relevo, para além da intensidade da vinculação do poder público, também na esfera das relações entre particulares, pois quando em causa uma violação da dignidade da pessoa humana advoga-se (embora se trate de entendimento não unânime) que até mesmo os atores privados estarão diretamente vinculados pelos direitos de igualdade, especialmente na sua dimensão negativa (defensiva), operando como proibições de discriminação.⁷⁴

Em semelhante sentido, Francisco Balaguer Callejón enfatiza a impossibilidade de se tolerar a reprodução de preconceitos e desrespeito às regras constitucionais de igualdade por meio algoritmos:

1.2. Algoritmos y derechos

Los problemas que plantean los algoritmos al sistema constitucional de derechos se extienden por una amplia relación de materias que afectan a un gran número de derechos constitucionales, ya sea mediante la utilización de los algoritmos ya sea mediante su propio diseño. En lo que se refiere a este último, se plantea de manera constante la cuestión de los sesgos y de su incidencia sobre los principios constitucionales que están orientados a promover la igualdad y a proteger a los grupos más vulnerables de la sociedad.

El argumento de que los sesgos son un reflejo de la realidad y que el diseño del algoritmo se limita a ajustarse a esa realidad carece de consistencia constitucional y resulta expresivo de la desatención permanente de los valores constitucionales por parte de las compañías tecnológicas. Esos sesgos son tan inaceptables como la realidad que reflejan. [...]

El choque con la constitución es inevitable porque en el ordenamiento constitucional de los países democráticos, de manera expresa o implícita (a través, por ejemplo, de la formulación de las políticas públicas) existe un mandato de promoción de la igualdad y de prohibición de la discriminación. La constitución no es neutral frente a la realidad, sino que proporciona en mayor o

⁷⁴ SARLET, Ingo Wolfgang. **Dignidade (da pessoa) humana e direitos fundamentais na Constituição Federal de 1988**. 10. ed. rev. atual. e ampl. Porto Alegre: Livraria do Advogado, 2019. p. 132.

*menor medida, la base cultural y normativa para promover, por ejemplo, la igualdad entre mujeres y hombres. El diseño de algoritmos con sesgos que normalicen y perpetúen la desigualdad de género es, por tanto, contrario a la constitución y no debe ser permitida por el ordenamiento jurídico.*⁷⁵

Como se vê, a questão ora discutida é relevante e produz efeitos diretos na vida das pessoas, eis que pode restringir acesso a oportunidades e direitos. Estes motivos levaram o já referido substitutivo aos projetos de lei que objetivam regular a inteligência artificial no Brasil a dedicar uma seção exclusiva para a discriminação algorítmica (Seção IV - Do direito à não-discriminação e à correção de vieses discriminatórios diretos, indiretos, ilegais ou abusivos).⁷⁶ Além disso, em seu artigo 4º, incisos VI e VII, a redação proposta define o que a lei entenderá por discriminação, no âmbito da inteligência artificial:

VI – discriminação: qualquer distinção, exclusão, restrição ou preferência, em qualquer área da vida pública ou privada, cujo propósito ou efeito seja anular ou restringir o reconhecimento, gozo ou exercício, em condições de igualdade, de um ou mais direitos ou liberdades previstos no ordenamento jurídico, em razão

⁷⁵ Em tradução livre: 1.2 Algoritmos e direitos - Os problemas colocados pelos algoritmos ao sistema constitucional de direitos estendem-se por uma ampla gama de assuntos que afetam um grande número de direitos constitucionais, seja através do uso de algoritmos ou através de seu desenho. Com relação a estes últimos, a questão dos preconceitos e seu impacto sobre os princípios constitucionais que visam promover a igualdade e proteger os grupos mais vulneráveis da sociedade é constantemente levantada.

O argumento de que os enviesamentos são um reflexo da realidade e que o desenho do algoritmo está meramente em conformidade com essa realidade carece de consistência constitucional e é indicativo do contínuo desrespeito aos valores constitucionais por parte das empresas de tecnologia. Tais preconceitos são tão inaceitáveis quanto a realidade que eles refletem.
[...].

O choque com a Constituição é inevitável porque na ordem constitucional dos países democráticos, expressa ou implicitamente (através, por exemplo, da formulação de políticas públicas), existe um mandato para promover a igualdade e proibir a discriminação. A constituição não é neutra diante da realidade, mas fornece, em maior ou menor grau, a base cultural e normativa para promover, por exemplo, a igualdade entre mulheres e homens. O desenho de algoritmos tendenciosos que normalizam e perpetuam a desigualdade de gênero é, portanto, contrário à constituição e não deve ser permitido pelo sistema legal. CALLEJÓN, Francisco Balaguer.

La constitución del algoritmo. Fundación Manuel Giménez Abad de Estudios Parlamentarios y del Estado Autonómico. Zaragoza, 2022. p. 40-41.

⁷⁶ CUEVA, Ricardo Villas Bôas et al. **Relatório final da Comissão de Juristas responsável por subsidiar elaboração de substitutivo sobre Inteligência Artificial no Brasil.** Brasília: Coordenação de Comissões Especiais, Temporárias e Parlamentares de Inquérito, 2022. Disponível em: <https://legis.senado.leg.br/sdleg-getter/documento?dm=9221643&ts=1671480646036>. Acesso em: 04 jan. 2023. p. 26.

de características pessoais como origem geográfica, raça, cor ou etnia, gênero, orientação sexual, classe socioeconômica, idade, deficiência, religião ou opiniões políticas.

VII – discriminação indireta: discriminação que ocorre quando normativa, prática ou critério aparentemente neutro tem a capacidade de acarretar desvantagem para pessoas pertencentes a grupo específico, ou as coloquem em desvantagem, a menos que essa normativa, prática ou critério tenha algum objetivo ou justificativa razoável e legítima à luz do direito à igualdade e dos demais direitos fundamentais.⁷⁷

Esclarecido o que se entende por discriminação algorítmica e como ela pode afetar direitos e garantias fundamentais, cabe compreender como ela ocorre e, mais adiante, como instrumentos legais já existentes em nosso ordenamento podem incentivar a implementação de cautelas para sua mitigação.

2.7 Como ocorre a discriminação algorítmica

Cathy O'Neil, no livro *Algoritmos de destruição em massa*⁷⁸, relata exemplos cometidos por algoritmos que reforçam a desigualdade social, reduzem oportunidades e, por isso, faz um alerta: tais situações tendem a provocar um *looping* eterno, cristalizando a condição das vítimas dos algoritmos, como se tratará a seguir. Demonstra em casos práticos que muitos dos modelos programam preconceitos, equívocos e vieses humanos, sem fornecer nenhum tipo de explicação para as partes envolvidas.⁷⁹

Para esclarecer, O'Neil conta como um algoritmo criado para prever a probabilidade de reincidência de réus foi racista: ao identificar que sentenças condenatórias eram mais severas quando o réu era afrodescendente, o judiciário americano investiu em um programa que calculava o potencial de reincidência dos presos e o utilizava para estabelecer sentenças e fianças. A adoção de modelos de

⁷⁷ CUEVA, *loc. cit.*

⁷⁸ O'NEIL, Cathy. **Algoritmos de destruição em massa**: como o big data aumenta a desigualdade e ameaça à democracia. Tradução Rafael Abraham. 1. ed. Santo André/SP: Rua do Sabão, 2020.

⁷⁹ *Ibidem*, p. 6.

risco computadorizados, alimentados por dados que não analisavam critérios como “cor da pele”, levaram as autoridades a acreditar que seriam proferidas sentenças mais justas, despidas de preconceito.

O *software* mencionado por O’Neil recebeu o nome de COMPAS⁸⁰ e ficou conhecido no julgamento do caso do réu Loomis vs. Wisconsin⁸¹. Loomis recorreu de sua punição, alegando que ela se baseou em um algoritmo cuja metodologia não foi divulgada ao réu e nem mesmo ao tribunal, por tratar-se de segredo comercial de propriedade da empresa fabricante. Apesar de não ter conseguido reverter judicialmente a decisão, o julgamento serviu para alertar os juízes de não se basearem exclusivamente nas indicações do programa, recomendando que os magistrados devem atuar na verificação de viés da ferramenta criada para corrigir seus próprios vieses.

Em que pese o resultado do julgamento não ter propiciado a comprovação do viés racista do algoritmo, uma pesquisa realizada pela ProPublica⁸² denunciou que o COMPAS era tendencioso contra negros.

Mesmo sem saber a cor da pele dos investigados, o algoritmo atribuía maior pontuação de risco para pessoas de pele escura, porque, dentre as informações mapeadas pelo algoritmo, estavam o endereço dos réus, suas relações com pessoas envolvidas em crimes, álcool e drogas – e, nos Estados Unidos da América (EUA), estas situações são mais presentes em bairros predominantemente habitados por negros e latinos em situação de vulnerabilidade socioeconômica, cujo policiamento mais rígido contribui para que um percentual considerável de seus habitantes possua antecedentes criminais em sua ficha ou em ficha de pessoas próximas, além de muitos deles já terem tido algum contato com drogas ou álcool.

⁸⁰ Sigla para *Correctional Offender Management Profiling for Alternative Sanctions*. Para mais informações, examinar: FERRARI, Isabela *et al.* **Justiça digital**. 1. ed. São Paulo: Thomson Reuters Brasil, 2020. p. 86.

⁸¹ STATE v. Loomis: Wisconsin Supreme Court requires warning before use of algorithmic risk assessments in sentencing. **130 Harvard Law Review**, 1530. Mar, 2017. Disponível em: <https://harvardlawreview.org/2017/03/state-v-loomis/>. Acesso em: 15 mar. 2022.

⁸² ANGWIN, Julia *et al.* Machine Bias: there's software used across the country to predict future criminals. And it's biased against blacks. **ProPublica**, 23 maio 2016. Disponível em: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. Acesso em: 15 mar. 2022.

Assim, mesmo sem saber a cor da pele, as características anteriormente mencionadas – contidas no questionário que embasa o cálculo da pontuação – são uma forma alternativa de segregar o mesmo grupo de pessoas.

Para reforçar o argumento, O'Neil demonstra o resultado de um estudo realizado em 2013 pela Associação Nova-Iorquina para Liberdades Civis que apontou que homens negros e latinos respondiam por 40,6% das paradas policiais, ainda que representassem apenas 4,7% da população da cidade. Mais de 90% destes revistados eram inocentes.⁸³

Por estarem constantemente sob a mira da desconfiança policial, tal classe possui respostas menos favoráveis para o questionário que projetará sua futura punição, acusando aqueles que a ela pertencem de ter mais chances de cometer crimes em razão de sua condição física e socioeconômica.

Esta prática influencia também nas estratégias de policiamento preditivo (que leva em consideração dados de crimes passados para tentar prever quais locais ou perfis de pessoas que possuem tendência para criminalidade), que está cada vez mais presente em cidades do Brasil e do mundo, apesar de seu potencial altamente discriminatório.

Assim, se houver altos índices de crimes cometidos por homens negros, o algoritmo tenderá a acreditar que um homem negro tem mais chances de cometer crimes do que mulheres ou que homens brancos.

Todavia, sabe-se que não é a cor da pele da pessoa que a condicionará ao cometimento de crimes, mas devido à opacidade sobre o aprendizado dos algoritmos e seus *proxies*, situações como esta podem ocorrer, ainda que o critério cor da pele não seja uma característica programada para ser levada em consideração – como no caso já citado *Loomis vs. Wisconsin*.

Além disso, a tendência é que esta situação acarrete uma redundância: policiamento ostensivo em determinadas áreas e grupos de suspeitos poderá levar a uma maior constatação de crimes nestes focos, que farão com que os bancos de dados sejam alimentados com estes novos delitos, reforçando a crença sobre o risco

⁸³ O'NEIL, Cathy. **Algoritmos de destruição em massa**: como o big data aumenta a desigualdade e ameaça à democracia. Tradução Rafael Abraham. 1. ed. Santo André/SP: Rua do Sabão, 2020. p. 29.

ser maior nestes locais/grupos e, conseqüentemente, ensejando mais vigilância sobre eles.⁸⁴

As conclusões demonstradas pela ProPublica⁸⁵ sobre os julgamentos induzidos pelo *software* COMPAS demonstram que utilizar tais preconceitos históricos como premissas para análises preditivas sobre a viabilidade de o criminoso reincidir são alarmantes, na medida em que, além de cometer injustiças, não se prestam ao seu objetivo final, que é impedir o cometimento de novos crimes, como demonstrado no Quadro 3 abaixo:

Quadro 3 – A previsão falha de forma diferente para réus negros

A previsão falha de forma diferente para réus negros		
	BRANCO	AFRO-AMERICANO
Rotulado como maior risco, mas não voltou a ofender	23,5%	44,9%
Rotulado como menor risco, mas reincidiu	47,7%	28,0%

No geral, a ferramenta de avaliação da Northpointe prevê corretamente a reincidência em 61% das vezes. Mas os negros são quase duas vezes mais propensos do que os brancos a serem rotulados como de maior risco, mas não de fato reincidirem. Comete o erro oposto entre os brancos: eles são muito mais propensos do que os negros a serem rotulados de menor risco, mas acabam cometendo outros crimes. (Fonte: análise ProPublica de dados de Broward County, Flórida.)

Fonte: ProPublica.⁸⁶

O Quadro 3 acima revela a preocupação de O'Neil⁸⁷ intrínseca aos modelos matemáticos, tendo em vista que eles se baseiam (i) em dados registrados do passado (desconsiderando o que não foi datificado) e (ii) na premissa/probabilidade de que o passado deve se repetir no futuro.

⁸⁴ KUNICHOFF, Yana; SIER, Patrick. The contradictions of Chicago police's secretive list. **Chicago Magazine**, 21 ago. 2017. Disponível em: <https://www.chicagomag.com/city-life/august-2017/chicago-police-strategic-subject-list/>. Acesso em: 15 mar. 2022.

⁸⁵ ANGWIN, Julia *et al.* Machine Bias: there's software used across the country to predict future criminals. And it's biased against blacks. **ProPublica**, 23 maio 2016. Disponível em: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. Acesso em: 15 mar. 2022.

⁸⁶ *Ibidem*.

⁸⁷ O'NEIL, Cathy. **Algoritmos de destruição em massa**: como o big data aumenta a desigualdade e ameaça à democracia. Tradução Rafael Abraham. 1. ed. Santo André/SP: Rua do Sabão, 2020. p. 29.

Para reforçar a ideia acima, cabe ilustrar com um exemplo prático descrito por Isabela Ferrari *et al.*:

Imagine, por exemplo que uma determinada empresa deseje contratar um funcionário que tenha chances de, no futuro, se tornar CEO, e que queira fazê-lo, como é corrente, empregando um software que pré-seleciona pessoas com o perfil adequado.

Como um software desse tipo seria treinado? Usando dados referentes às pessoas que são CEOs, atualmente. E qual o perfil dos CEOs na maior parte das empresas? Homens brancos de 60 anos. Se o algoritmo for treinado a partir dessa informação, que é fidedigna, é provável que selecione, majoritariamente, homens brancos.

Ou seja, o programa vai tender a rechaçar, por exemplo, a contratação de mulheres, porque os dados que ele recebe simplesmente não compreendem um número significativo de mulheres e não porque elas sejam menos capazes.⁸⁸

Vê-se, portanto, que a IA tem a capacidade de analisar e encontrar características comuns em grandes bancos de dados, e repeti-los quando encontra padrões de situações semelhantes, mas ainda não tem a capacidade de aprender sozinha a fazer o juízo de valor das características que são positivas ou negativas de serem reproduzidas.

A pesquisadora do Massachusetts Institute of Technology (MIT), Joy Buolamwini, também desvendou um preconceito racial e de gênero em serviços de inteligência artificial de empresas como a Microsoft, IBM e Amazon. Realizando seu trabalho de conclusão, Buolamwini desenvolveu uma aplicação que não reconhecia sua própria face, de pele escura.⁸⁹

Ao investigar o motivo desta falha, constatou que treinou seu algoritmo com base em um banco de dados abertos contendo milhões de imagens de rostos humanos. Todavia, o banco não continha uma amostra de dados significativa para rostos femininos, especialmente rostos femininos e negros, o que impediu que o

⁸⁸ FERRARI, Isabela *et al.* **Justiça digital**. 1. ed. São Paulo: Thomsom Reuters Brasil, 2020. p. 90.

⁸⁹ BUOLAMWINI, Joy; GEBRU, Tinmit. "Gender shades: intersectional accuracy disparities in commercial gender classification." **Proceedings of Machine Learning Research**, v. 81, n. 1, p. 1–15, 2018. Conference on Fairness, Accountability, and Transparency.

algoritmo pudesse aprender a reconhecê-los com precisão. Ou seja, neste caso foi fornecida apenas uma parcela de realidade – um recorte que não ensina sobre a totalidade do mundo real.

Aprofundando sua pesquisa, Buolamwini constatou que esta limitação era comum em diversos aplicativos de reconhecimento facial, incluindo os desenvolvidos/utilizados pelas grandes empresas anteriormente mencionadas. Seu trabalho foi documentado em vídeo e está disponível no Netflix com o título '*Coded Bias*'.⁹⁰

Outras grandes empresas como Google e Facebook também possuem ferramentas de reconhecimento facial amplamente utilizadas, porém são inúmeras as vezes que indicam uma pessoa totalmente diferente da que está na foto ou no vídeo. Os algoritmos empregados atualmente podem ser muito bons, mas não são infalíveis.

Ainda que a tecnologia hoje tenha alto grau de acurácia, no Rio de Janeiro, uma mulher foi levada para a delegacia após ser identificada como foragida por uma das câmeras de reconhecimento facial. Ao chegar lá, perceberam que se tratava de um engano – até mesmo porque a suposta foragida já estava presa. Ou seja, além da imprecisão do algoritmo utilizado, o banco de dados estava desatualizado.⁹¹ Em outro caso, um homem de 25 anos com necessidades especiais teve armas apontadas para sua cabeça até que os policiais percebessem que se tratava de um engano da tecnologia.⁹²

A IA utilizada por buscadores da *web* para ranquear resultados também pode ser discriminatória. Nem mesmo o Google passou ileso de críticas ao reproduzir

⁹⁰ Estes são os dados do documentário mencionado: CODED BIAS. *In*: WIKIPÉDIA: a enciclopédia livre. [San Francisco: Wikimedia Foundation, 2021]. Disponível em: https://en.wikipedia.org/wiki/Coded_Bias. Acesso em: 23 jan. 2023.

⁹¹ LAVADO, Thiago. Aumento do uso de reconhecimento facial pelo poder público no Brasil levanta debate sobre limites da tecnologia. **G1**, 21 fev. 2020. Disponível em: <https://g1.globo.com/economia/tecnologia/noticia/2020/02/21/aumento-do-uso-de-reconhecimento-facial-pelo-poder-publico-no-brasil-levanta-debate-sobre-limites-da-tecnologia.ghtml>. Acesso em: 15 mar. 2022.

⁹² PALMA, Amanda; PACHECO, Clarissa. Entenda como funciona o reconhecimento facial que ajudou a prender mais de 100 na BA - Câmeras estão espalhadas em várias partes de Salvador e imagens são enviadas à central da SSP. **Correio 24 horas**. Salvador, 05 jan. 2020. Disponível em: <https://www.correio24horas.com.br/noticia/nid/entenda-como-funciona-o-reconhecimento-facial-que-ajudou-a-prender-mais-de-100-na-ba>. Acesso em: 15 mar. 2022.

conceitos totalmente preconceituosos quando relacionados a mulheres.⁹³ Após denúncia de duas artistas brasileiras, a empresa precisou se retratar publicamente por conceituar os termos “mulher solteira” como “prostituta, meretriz” e “patroa” como “a mulher do patrão. Dona de casa”.

Em contrapartida, quando os mesmos termos eram pesquisados no substantivo masculino, os conceitos apareciam como “aquele que não casou” e “proprietário ou chefe de um estabelecimento privado comercial”, respectivamente. Em comunicado, a empresa explicou o que ocasionou sua postura machista:

Não editamos nem removemos as definições fornecidas pelos nossos parceiros que são os especialistas em idiomas. No caso dos significados das palavras “patroa” e “mulher-solteira”, a Oxford Languages, nossa parceira que trabalha com uma das editoras mais tradicionais de dicionários do Brasil, determinou que ambas definições não refletem mais o uso moderno da língua portuguesa falada pelos brasileiros e não são usadas o suficiente para serem incluídas nos resultados de significados.⁹⁴

Veja-se que, no caso descrito acima, bem como nos anteriormente mencionados, não houve intencionalidade nas discriminações reproduzidas, até porque muitas vezes ela ocorre por espelhar dados relativos a fatos históricos – muitos já superados ou ainda em vias de superação.

Até mesmo o ChatGPT (sigla de *Generative Pre-training Transformer*) desenvolvido pela OpenAI em dezembro de 2022 não passou ileso. O aplicativo, visto por muitos como a ferramenta que irá substituir o buscador Google, funciona como um *chat* em que o usuário pode fazer perguntas sobre qualquer tema.⁹⁵ Em vez de

⁹³ MIOTTO, Rafael. Significados de 'patroa' e 'mulher-solteira' mudam no Google depois de críticas de Anitta e Luísa Sonza. **G1**, 18 set. 2020. Disponível em: <https://g1.globo.com/economia/tecnologia/noticia/2020/09/18/significados-de-patroa-e-mulher-solteira-mudam-no-google-depois-de-criticas-de-anitta-e-luisa-sonza.ghtml> Acesso em: 15 mar. 2022.

⁹⁴ MIOTTO, Rafael. Significados de 'patroa' e 'mulher-solteira' mudam no Google depois de críticas de Anitta e Luísa Sonza. **G1**, 18 set. 2020. Disponível em: <https://g1.globo.com/economia/tecnologia/noticia/2020/09/18/significados-de-patroa-e-mulher-solteira-mudam-no-google-depois-de-criticas-de-anitta-e-luisa-sonza.ghtml> Acesso em: 15 mar. 2022.

⁹⁵ Por exemplo, para fins de teste, foi solicitado de forma simples ao Chat GPT que ele explicasse sobre vieses algorítmicos e quais as formas de combatê-los. O aplicativo forneceu a seguinte resposta (em

indexar sites que contenham a informação buscada, ele responde ao usuário com um texto explicativo sobre o tema questionado. Ele também executa comandos, como, por exemplo, desenvolver um texto, uma poesia e até códigos de programação.

Apesar de todos os esforços da OpenAI⁹⁶ para desenvolver uma aplicação que não reproduza preconceitos, entre outras ofensas a direitos, diversas respostas

tradução livre): O viés do algoritmo é a tendência de um algoritmo de produzir resultados que são injustamente tendenciosos ou discriminatórios em relação a certos grupos de pessoas. Isso pode ocorrer de várias maneiras, como por meio do uso de conjuntos de dados falhos, da incorporação de vieses humanos subjetivos ao algoritmo ou do uso de métricas de avaliação inadequadas ou inapropriadas. Um exemplo comum de viés de algoritmo é o uso de software de reconhecimento facial que tem maior probabilidade de identificar erroneamente indivíduos de certas raças ou etnias. Isso geralmente se deve ao fato de que os conjuntos de dados usados para treinar o algoritmo não são representativos de toda a população e, como resultado, é mais provável que o algoritmo cometa erros quando confrontado com indivíduos que não se encaixam no perfil "típico". Outro exemplo de viés de algoritmo pode ocorrer em práticas de contratação, em que determinados algoritmos podem favorecer determinados candidatos em detrimento de outros com base em sexo, idade ou outras características demográficas. Isso pode levar à discriminação e à falta de diversidade no local de trabalho. Para mitigar os efeitos do viés do algoritmo, é importante avaliar cuidadosamente os conjuntos de dados e as métricas de avaliação usadas no desenvolvimento do algoritmo e buscar e incorporar ativamente diversas perspectivas no processo de design. Também é importante monitorar e testar regularmente o algoritmo quanto a viés e implementar estratégias para corrigir quaisquer vieses identificados. O viés do algoritmo pode ser um problema complexo e cheio de nuances, e não há respostas fáceis ou soluções rápidas. No entanto, existem algumas etapas que organizações e indivíduos podem seguir para ajudar a reduzir a probabilidade de viés nos algoritmos:

Esteja ciente do potencial de viés: Este é o primeiro passo para lidar com o viés nos algoritmos. Reconhecer que o viés pode existir e ser um problema é o primeiro passo para evitá-lo;

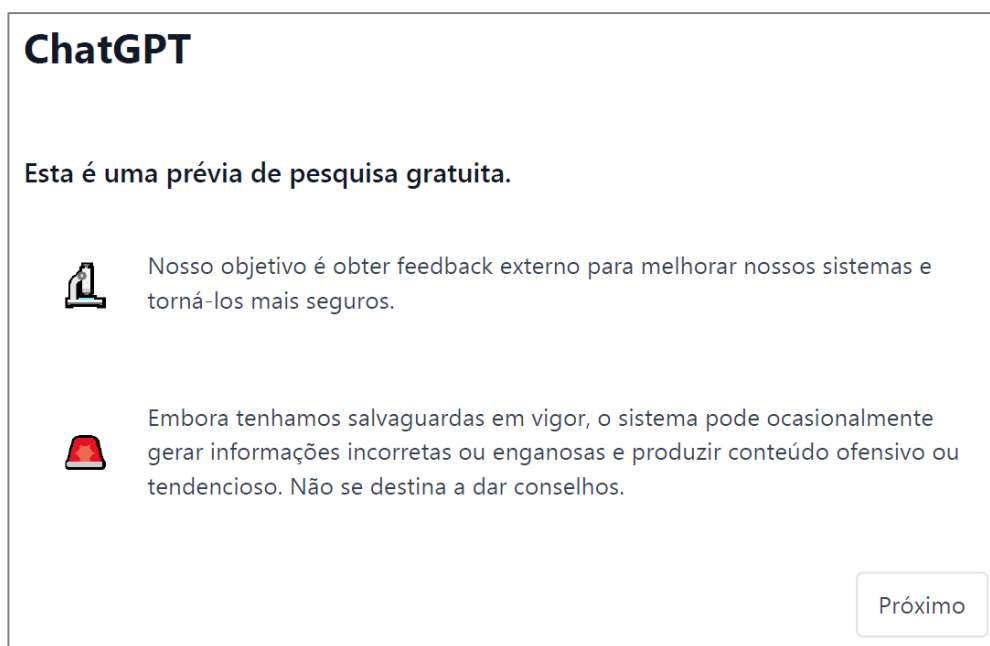
- 1) Examine e limpe seus dados: os dados usados para treinar e testar algoritmos podem conter vieses. É importante examinar os dados e remover quaisquer vieses que possam estar presentes;
- 2) Use métricas de avaliação justas: as métricas usadas para avaliar o desempenho de um algoritmo também podem ser tendenciosas. É importante usar métricas de avaliação justas que não favoreçam um grupo em detrimento de outro;
- 3) Monitorar e testar algoritmos: é importante testar e monitorar regularmente os algoritmos para garantir que eles estejam funcionando conforme o esperado e não exibam nenhum viés;
- 4) Use diversos conjuntos de treinamento: usar um conjunto de treinamento diversificado pode ajudar a reduzir a probabilidade de viés nos algoritmos;
- 5) Aumentar a transparência: Tornar o processo e a tomada de decisão dos algoritmos mais transparentes pode ajudar a identificar e lidar com quaisquer vieses que possam existir;
- 6) Considere os impactos potenciais: é importante considerar os impactos potenciais de um algoritmo em diferentes grupos de pessoas e tomar medidas para mitigar quaisquer impactos negativos.

No geral, é importante ser proativo e vigilante ao abordar a questão do viés nos algoritmos. *In: OPENAI, 2023. Disponível em: <https://openai.com/about/>. Acesso em: 09 jan. 2023.*

⁹⁶ A entidade se autodescreve como: *OpenAI is an AI research and deployment company. Our mission is to ensure that artificial general intelligence benefits all of humanity. OpenAI's mission is to ensure that artificial general intelligence (AGI)—by which we mean highly autonomous systems that outperform humans at most economically valuable work—benefits all of humanity. We will attempt to directly build safe and beneficial AGI, but will also consider our mission fulfilled if our work aids others to achieve this outcome. OPENAI is an AI research and deployment company. Our mission is to ensure that artificial general intelligence benefits all of humanity.* **Em tradução livre:** Nossa missão é garantir que a inteligência artificial geral – sistemas de IA geralmente mais inteligentes que os humanos – beneficie toda a humanidade. Estamos construindo uma AGI segura e benéfica, mas também

xenofóbicas já foram reportadas⁹⁷, de forma que a instituição faz um alerta aos usuários (Figura 3):

Figura 3 – Avisos dados antes de utilizar o ChatGPT



Fonte: OpenAI.⁹⁸

Neste sentido, Cintia Freitas e Jean Paul Barddal explicam o motivo de estes fenômenos ocorrerem, mesmo que se adotem determinadas cautelas:

Na sociedade contemporânea, indivíduos e empresas produzem enormes quantidades de dados. Desde leituras de sensores em telefones celulares e gadgets até postagens nas mídias sociais, dados são produzidos de forma contínua e abundante. Importante perceber que tais dados representam as atividades e comportamentos diários de cada pessoa e, portanto, podem ser utilizados em análises preditivas. [...] Neste contexto, a análise preditiva é uma

consideraremos nossa missão cumprida se nosso trabalho ajudar outras pessoas a alcançar esse resultado. In: OPENAI, 2023. Disponível em: <https://openai.com/about/>. Acesso em: 09 jan. 2023.

⁹⁷ BIDDLE, Sam. The internet's new favorite AI proposes torturing Iranians and surveilling mosques. **The Intercept**. 08 dez. 2022. Disponível em: <https://theintercept.com/2022/12/08/openai-chatgpt-ai-bias-ethics/>. Acesso em: 09 jan. 2023.

⁹⁸ OPENAI, *op. cit.*

abordagem popular para obter informações e padrões sobre os dados e criar modelos preditivos. A análise preditiva visa aproveitar os dados do passado para obter informações em tempo real e prever eventos futuros.⁹⁹

Mas, será que os recortes e arranjos de realidades que estão sendo fornecidos para que os robôs aprendam correspondem ao futuro almejado? Se importantes decisões que influenciam no comportamento da nossa sociedade são tomadas pela IA, com base no que ela identifica e classifica sobre os dados captados a partir de coisas que já ocorreram, sem realmente “aprender” com os erros já cometidos, como a humanidade irá evoluir?

Utilizando como exemplo os pais e as escolas: ao ensinarem fatos históricos às crianças, as ensinam a não repetir indiscriminadamente todos os fatos do passado, para que possa haver melhores práticas de convívio social. Assim, talvez o futuro a ser ensinado para as máquinas, a depender do caso, devesse ser um modelo totalmente novo e diferente do que simplesmente reproduzir dados do passado.

É possível denotar, portanto, que, no atual momento tecnológico, ainda não há aprendizagem de máquina, mas sim uma repetição de máquina, que apenas replica em dados futuros os mesmos padrões que identificou em dados anteriores. Esta simples substituição de um termo não busca menosprezar o avanço tecnológico ou desrespeitar o termo técnico amplamente utilizado, mas teria o condão de auxiliar os usuários a entender do que realmente se trata, para não confiarem cegamente nas sugestões e tomadas de decisão realizadas por IA.

Vê-se que, na prática, a IA não está aprendendo (no sentido de fazer juízo crítico de valor) a partir do que analisou, só está repetindo comandos a partir das orientações que recebeu e das correlações que encontrou nos dados que estavam disponíveis para sua pesquisa.

Neste mesmo sentido, a terminologia “análise preditiva” induz à ideia de que a tecnologia, altamente avançada, tem a capacidade de prever o futuro, quase como se pudesse adivinhá-lo. Contudo, como visto em tópico anterior, na verdade,

⁹⁹ FREITAS, Cinthia Obladen de Almendra; BARDDAL, Jean Paul. Análise preditiva e decisões judiciais: controvérsia ou realidade? **Revista Democracia Digital e Governo Eletrônico**, Florianópolis, v. 1, n. 18, p. 107-126, 2019.

trata-se da indicação de um resultado provável, com base em regras de probabilidade matemática e/ou estatística. Possivelmente o termo análise “probabilística” ou “estatística” inspiraria maiores cautelas antes de os usuários acatarem as sugestões das máquinas.

Contudo, fortes nesta confiança cega que a tecnologia pode provocar, organizações delegam aos julgamentos algorítmicos decisões a respeito de diferentes dimensões da autonomia privada dos cidadãos, na medida em que são estes sistemas que resolvem quem ingressará em universidades ou terá acesso a empregos, cargos e oportunidades profissionais.¹⁰⁰

Ana Frazão ressalta que em alguns casos é a própria vida humana que será julgada, como por exemplo nos comandos dados a carros autônomos, que precisarão decidir quais vidas serão poupadas e quais não, em caso de acidentes inevitáveis. E tudo isso é feito “sem as devidas considerações éticas e jurídicas que deveriam orientar a sua adoção.”¹⁰¹

Na verdade, a máquina sequer sabe julgar que determinado fato é um acerto ou um erro (a ser descartado como um modelo a seguir), a menos que seja programada para isto a partir do critério classificatório eleito por um humano (que também pode ser um critério equivocado ou tendencioso, mesmo que não intencionalmente). Sendo assim, a etapa seguinte desta pesquisa se dedicará a compreender as possíveis causas da discriminação algorítmica, no intuito de, a partir delas, verificar quais as ferramentas possíveis para sua mitigação.

2.8 Possíveis causas

A partir da abordagem a respeito da discriminação algorítmica, Duarte e Negócio buscaram compreender como ocorre a perpetuação de desigualdade por meio de algoritmos e pela inteligência artificial.

¹⁰⁰ FRAZÃO, Ana. Discriminação algorítmica: compreendendo o que são os julgamentos algorítmicos e o seu alcance na atualidade. Parte I. **JOTA**. 16 jun. 2021. Disponível em: <https://www.jota.info/opiniao-e-analise/colunas/constituicao-empresa-e-mercado/discriminacao-algoritmica-16062021>. Acesso em: 18 nov. 2021.

¹⁰¹ *Ibidem*.

Assim, descrevem os autores que, diferentemente das automações não inteligentes, cujos efeitos discriminatórios são oriundos de programação intencional ou negligência, a inteligência artificial possui mecanismos que podem acidentalmente ocasionar processos discriminatórios. Esclarecem que tal fato pode ocorrer em razão da necessidade de reduzir análises complexas à linguagem computacional para execução pelo algoritmo. Essa simplificação de análises ou problemas, chamada pelos autores de “tradução”, é exemplificada no artigo da seguinte maneira:

Se o desenvolvedor rotula um conjunto de dados, os quais são usados para treinar o modelo, como “bom pagador” e outro como “mau pagador” e faz isso com base em critérios próprios, ou seja, naquilo que ele acredita ser a definição mais adequada de “bom/mau pagador”, ele estará expondo ao modelo um conjunto de dados cujos rótulos refletem uma perspectiva específica e subjetiva sobre um fenômeno maior e o sistema irá inferir desse conjunto de dados um certo padrão, o qual será aplicado por outros dados distintos daqueles do dataset training. Essas novas classes criadas pelo desenvolvedor apresentam alguns pontos cegos (O’Neil, 2016) em razão da limitação epistêmica – ou seja, a definição do que é bom adotada refletirá apenas um aspecto dentro de um amplo espectro de possibilidades – os quais poderão ser irrelevantes ou, ao contrário, gera efeitos nefastos.¹⁰²

Outro problema mapeado pelos autores como causador de possíveis discriminações algorítmicas refere-se às bases de dados utilizadas para treinamento das aplicações: dados incompletos (com sub ou super-representação de determinados grupos) ou enviesados podem fazer com que algoritmos aprendam padrões tendenciosos e discriminatórios. Explicam “em poucas palavras, os dados refletem situações passadas e, ao serem incorporados nos algoritmos de tomada de

¹⁰² DUARTE, Alan; NEGÓCIO, Ramon de Vasconcelos. Todos são iguais perante o algoritmo? uma resposta cultural do Direito à discriminação algorítmica. **Revista Direito Público**, Brasília, v. 18, n. 100, 2022. DOI: 10.11117/rdp.v18i100.5869. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/5869>. Acesso em: 4 ago. 2022. p. 229.

decisão, definirão comportamentos futuros.”¹⁰³ Ainda, observam a necessidade de atenção na fase de coleta de dados, os quais precisam ser capazes de representar adequadamente os diferentes grupos sociais.

Também ressaltam a existência de grupos à margem do *big data*, que, por não possuírem acesso à tecnologia, não estão datificados e, com isso, não serão representados nos bancos de dados que treinam o *machine learning*, fato que pode distorcer algumas decisões.

Por último, os pesquisadores atentam para o uso de *proxies* como um outro possível causador de discriminações algorítmicas. Conceituam *proxy* como uma variável que, a princípio, não tem vínculo com o problema – contudo, acaba por ser utilizada como uma correlação indicativa da informação que se precisaria obter. Para ilustrar, reproduz-se o trecho a seguir:

Utiliza-se, para exemplificar esse ponto, o caso em que a contratação de pessoas é feita por algoritmos de aprendizado de máquina e um dos critérios utilizados para selecionar o melhor candidato seja o nível de escolaridade e a instituição de ensino que o candidato frequentou. Caso as melhores instituições (assim consideradas pelo desenvolvedor do sistema) possuam um elevado número de estudantes brancos e homens, então ocorrerá uma codificação redundante, ou seja, a cor e o gênero serão codificados dentro do critério escolaridade.^{104.105}

¹⁰³ DUARTE; NEGÓCIO, *loc. cit.*

¹⁰⁴ DUARTE, Alan; NEGÓCIO, Ramon de Vasconcelos. Todos são iguais perante o algoritmo? uma resposta cultural do Direito à discriminação algorítmica. **Revista Direito Público**, Brasília, v. 18, n. 100, 2022. DOI: 10.11117/rdp.v18i100.5869. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/5869>. Acesso em: 4 ago. 2022. p. 231.

¹⁰⁵ No caso exemplificado, para estabelecer um conceito de melhor candidato, definiu-se como *proxy* a instituição de ensino que o candidato se graduou – critério que indica um dado ou uma característica. Todavia, tal dado não indica se o candidato tem um conjunto de características mais adequadas para a vaga, pois sabe-se que não é a instituição que a pessoa estudou que define se ela é a melhor candidata ou não. Como o objetivo buscado, ou seja, “o melhor candidato”, é um conceito subjetivo, há necessidade de correlacionar ele com um critério binário, possível de ser mapeado por dados objetivos, permitindo que os algoritmos tenham um parâmetro para classificar quais candidatos são os melhores (aqui definidos como aqueles que estudaram na instituição X), dentro de um mar de currículos. Destaca-se que tal *proxy* pode acabar excluindo candidatos excelentes (e até melhores) por terem estudado em outras instituições não eleitas pelo critério - por isso é que o *proxy* pode ser discriminatório, eis que já exclui pessoas do acesso/oportunidade de uma entrevista, por exemplo.

Mendes e Mattiuzzo esclarecem que um dos principais objetivos dos algoritmos é analisar dados, extrair probabilidades e, por meio delas, indicar previsões. A crescente disponibilização de dados, captados em maior volume e variedade (*big datas*)¹⁰⁶, fez com que milhares de decisões fossem automatizadas, impactando a vida de todos. Quanto aos *big datas*, explicam que há três características que podem ter em um *big data*: (i) grande quantidade de dados e de informações coletadas; (ii) possível imprecisão dos dados disponíveis, em razão de seu grande volume e (iii) a busca de correlações em vez de causalidades entre os dados mapeados.

Com isso, destacam que a última característica é relevante para estudo da discriminação algorítmica, eis que por muito tempo a ciência funcionou com foco na busca de causalidades, por entender-se ser este o único modo de entender o porquê de os problemas ocorrerem e como identificar soluções a partir da constatação de sua causa. Contudo, no *big data*, a causalidade é substituída pelas correlações, que configuram uma "probabilidade de um evento ocorrer caso outro evento também se realize. É uma relação estatística entre tais acontecimentos. Ao invés de tentar assimilar os mecanismos internos de um fenômeno, as correlações nos permitem compreender o mundo por meio de proxies."¹⁰⁷

¹⁰⁶ "A função mais importante de Big Data é elaborar previsões baseadas em um grande número de dados e informações: desde desastres climáticos até crises econômicas, do surto de uma epidemia até o vencedor de um campeonato de esportes, do comportamento de um consumidor até a solvência dos clientes. Assim, as análises de Big Data podem ser utilizadas para elaborar prognósticos, tanto com relação à economia, à natureza ou à política, quanto sobre comportamento individual. No que se refere ao assunto aqui discutido, a predição do comportamento individual é de grande interesse, na medida em que gerar informação e conhecimento sobre o comportamento de uma pessoa a partir de dados pessoais oferece base para tomada de decisões. Uma análise de Big Data pode, portanto, afetar diretamente um indivíduo – e produzir resultados discriminatórios que impactem sua vida. De acordo com Mayer-Schönberger e Cukier, não há definição precisa para Big Data, mas o fenômeno pode ser caracterizado por três tendências. Em primeiro lugar, a quantidade de dados e informações coletadas. As análises de Big Data não apenas reúnem mais dados do que nunca, mas buscam reunir todos os dados e informações referentes a uma situação em particular, não apenas uma amostra deles – como colocam os autores, em Big Data, n = tudo." In: MENDES, Laura Schertel; MATTIUZZO, Marcela. Discriminação

algorítmica: conceito, fundamento legal e tipologia. **Revista Direito Público**, Brasília, v. 16, n. 90, 2019. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/3766>. Acesso em: 6 ago. 2022. p. 44.

¹⁰⁷ MENDES, Laura Schertel; MATTIUZZO, Marcela. Discriminação algorítmica: conceito, fundamento legal e tipologia. **Revista Direito Público**, Brasília, v. 16, n. 90, 2019. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/3766>. Acesso em: 6 ago. 2022. p. 44.

Segundo Mendes e Mattiuzzo, o uso de correlações com base em *proxies* tem sido cada vez mais utilizado com a disseminação da técnica chamada de *machine learning* (ML) ou aprendizagem de máquina, em que os próprios algoritmos se adaptam na medida que recebem mais dados e *feedbacks* dos resultados indicados. Em razão desta autoprogramação algorítmica, há uma obscuridade (chamada de caixas preta ou *black box*) que dificulta o entendimento do “caminho” que levou a determinada decisão. Exemplificando as afirmações, as autoras mencionam o caso COMPAS (já referido nesta pesquisa).

Comentam as autoras que as empresas têm comumente utilizado *machine learning* para recrutamento e seleção de currículos de candidatos a vagas de trabalho, fato que pode ser positivo, mas também perigoso:

Além disso, o uso de algoritmos tem o potencial de levar a resultados consistentes e a ajudar as empresas a economizarem – tendo em vista que os processos de contratação consomem muito tempo e dinheiro. De fato, os seres humanos são tendenciosos e é sabido que tais predisposições influenciam processos de recrutamento. Nesse sentido, o uso de algoritmos para selecionar candidatos pode minimizar discriminações. O problema, porém, é que, quanto mais recorremos a ferramentas como essas, mais difícil se torna para os candidatos “fora dos padrões” a entrada no mercado de trabalho e mais difícil é entender o que exatamente faz com que sua candidatura seja menos atrativa do que as outras.¹⁰⁸

Neste sentido, diagnosticam que a falta de transparência é uma preocupação que envolve o tema da discriminação algorítmica, pois a obscuridade impede que se avalie que algum resultado discriminatório foi efetivamente produzido, impedindo, também, que se trate o problema com o fito de que futuras discriminações sejam compelidas. Deste modo, podem acabar por reforçar preconceitos já existentes na sociedade.

¹⁰⁸ *Ibidem*, p. 46.

Ainda, as pesquisadoras categorizaram os quatro principais tipos de discriminação algorítmica identificados: (i) por erro estatístico, (ii) por generalização, (iii) pelo uso de informações sensíveis e (iv) limitadora do exercício de direitos.

A discriminação por erro estatístico é aquela que ocorre a partir de dados incorretamente coletados ou por erro de programação, prejudicando a análise estatística.

A discriminação por generalização decorre de uma classificação incorreta de um indivíduo a partir do grupo a que ele está inserido – como, por exemplo, quando se atribui baixo índice de adimplemento a um sujeito que reside num bairro humilde, sem considerar sua renda ou seu histórico individual. Frequentemente estas decisões discriminatórias são baseadas em informações que não estão diretamente vinculadas ao resultado que se busca (*proxies*), mas, estatisticamente, podem coincidir. Tal tipo põe em xeque a justiça individual e a desigualdade, pois enseja que “tratamentos diferenciados ocorram com base em características pessoais, na medida em que essas características são, de acordo com hipóteses estatísticas, um aspecto relevante para a tomada de decisões.”¹⁰⁹

A discriminação pelo uso de informações sensíveis pode ser estatisticamente correta, porém se vale de condições do indivíduo que já são objeto de preconceitos históricos, com base em dados que são inclusive protegidos legalmente. Tal categoria também pode ser endógena e reforçar padrões: “por tratarem de grupos historicamente discriminados, é um dos tipos mais perversos de discriminação, ao reforçar o tratamento discriminatório e automatizá-lo, tornando mais difícil para os membros de tais agrupamentos superarem determinada situação prejudicial.”¹¹⁰

E, por último, a discriminação limitadora do exercício de direitos também pode ser relevante e baseada em dados estatisticamente corretos, porém, ao ser a razão de afetar o exercício de determinado direito por um grupo, é discriminatória.

Para ilustrar, segue quadro-resumo (Quadro 4):

¹⁰⁹ MENDES, Laura Schertel; MATTIUZZO, Marcela. Discriminação algorítmica: conceito, fundamento legal e tipologia. **Revista Direito Público**, Brasília, v. 16, n. 90, 2019. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/3766>. Acesso em: 6 ago. 2022. p. 53.

¹¹⁰ *Ibidem*, p. 54.

Quadro 4 – Quadro-resumo discriminação algorítmica conforme Mendes e Mattiuzzo

Discriminação Algorítmica	
Tipo	Descrição
Por erro estatístico	Ocorre a partir de dados incorretamente coletados ou por erro de programação, prejudicando a análise estatística.
Por generalização	Decorre de uma classificação incorreta de um indivíduo a partir do grupo a que ele está inserido.
Pelo uso de informações sensíveis	Pode ser estatisticamente correta, porém se vale de condições do indivíduo que já são objeto de preconceitos históricos, com base em dados que são inclusive protegidos legalmente.
Limitadora do exercício de direitos	Também pode ser relevante e baseada em dados estatisticamente corretos, porém ao ser a razão de afetar o exercício de determinado direito por um grupo, é discriminatória.

Fonte: Elaborado pela autora com base em Mendes e Mattiuzzo.¹¹¹

As categorizações criadas exigem diferentes formas de mitigá-las, o que, segundo Mendes e Mattiuzzo, deve ser feito por meio de estratégias de governança algorítmica, as quais serão comentadas em capítulo específico.

Em semelhantes conclusões chegaram Doneda *et al.* Os autores descrevem os inúmeros relatos da literatura sobre os riscos de discriminação, mesmo que ela ocorra de forma involuntária e sem o conhecimento dos desenvolvedores. Segregam em dois grandes grupos as causas de tal discriminação, sendo a primeira os dados e, a segunda, estatística.

Com relação aos dados, apontam que o algoritmo será tão bom quanto os dados que o alimentam, o que quer dizer que, se lhe forem fornecidos dados eivados de preconceitos históricos para fins de treinamento, seus resultados refletirão estes mesmos preconceitos históricos.

¹¹¹ MENDES, Laura Schertel; MATTIUZZO, Marcela. Discriminação algorítmica: conceito, fundamento legal e tipologia. **Revista Direito Público**, Brasília, v. 16, n. 90, 2019. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/3766>. Acesso em: 6 ago. 2022. p. 54.

Já a discriminação estatística ocorre quando os indivíduos são classificados conforme características que possuem em comum com algum grupo, e essas características acabam sendo vinculadas a outros atributos que estatisticamente apresentem correlação. Neste tipo, a discriminação pode ocorrer por erro estatístico decorrente de dados incorretos ou de bases científicas frágeis.

Ainda mencionam a discriminação que pode ocorrer em razão de generalização dos indivíduos, pois, ao enquadrar as pessoas em rótulos em razão de uma ou mais características que possuem, deixa-se de analisar outros aspectos que podem ser relevantes para a tomada de decisão, especialmente quando o indivíduo é atípico ao grupo. É o que ocorre, por exemplo, quando se atribui baixa pontuação de crédito aos indivíduos de um determinado bairro considerado mais humilde. Ao eleger este critério (endereço) como base para a tomada de decisão (concessão ou não de crédito), tem-se a possibilidade de discriminar aqueles que possuem alta renda. Abaixo, segue o quadro-resumo destas categorias (Quadro 5):

Quadro 5 – Quadro-resumo discriminação algorítmica conforme Doneda et al.

Discriminação Algorítmica	
Tipo	Descrição
Dados	Ocorre se lhe forem fornecidos dados eivados de preconceitos históricos para fins de treinamento, seus resultados refletirão estes mesmos preconceitos históricos
Estatística	Ocorre quando os indivíduos são classificados conforme características que possuem em comum com algum grupo, e essas características acabam sendo vinculadas a outros atributos que estatisticamente apresentem correlação. Pode decorrer de: 1) dados incorretos 2) bases científicas frágeis
Generalização	Ocorre quando se enquadra as pessoas em rótulos em razão de uma ou mais características que possuem, sem analisar outros aspectos que podem ser relevantes para a tomada de decisão, especialmente quando o indivíduo é atípico ao grupo.

Fonte: Elaborado pela autora com base em Doneda et al.¹¹²

¹¹² DONEDA, Danilo Cesar Maganhoto et al. Considerações iniciais sobre inteligência artificial, ética e autonomia pessoal. **Pensar – Revista de Ciências Jurídicas**, Fortaleza, v. 23, n. 4, p. 1-17, out/dez. 2018. DOI: 10.5020/2317-2150.2018.8257. Disponível em: <https://ojs.unifor.br/rpen/article/view/8257>. Acesso em: 05 out. 2022.

Os autores constatam que os problemas de discriminação estão comumente vinculados a questões que envolvem *big data*¹¹³, por ser difícil assegurar a correção e atualização de todos os dados, ocasionando confusão e imprecisão em um método que aparenta ser objetivo, o que dificulta ainda mais a identificação do que afetou determinado direito.

Outro ponto de atenção que também foi indicado pelos autores é a falta de transparência do modo como determinada decisão foi tomada pelos algoritmos, gerando o que se chama de obscuridade ou "*black box*", que é a impossibilidade de terceiros averiguarem as fórmulas e *inputs* do algoritmo em razão de segredo industrial, e dificultando a identificação de prática discriminatória.

Assim como para Mendes e Mattiuzzo, para Doneda *et al.* também é objeto de preocupação a busca de correlações entre as situações mapeadas pelos dados em substituição à causalidade, para basear as previsões dadas pelas ferramentas de IA. Isso porque, muitas vezes, as correlações feitas entre os eventos serão totalmente aleatórias, sem qualquer vínculo de causa e efeito – o que pode ocasionar o que chamam de discriminação por associação ou discriminação por *proxy*:

Como visto acima, apesar dos benefícios das decisões automatizadas, como a maior agilidade no processamento de informações e a redução dos custos associados à gestão e à contratação de funcionários, tais decisões podem afetar significativamente os direitos fundamentais do indivíduo, em especial a sua autonomia, personalidade e igualdade. Embora a disciplina da proteção de dados pessoais por meio da regulação do fluxo de informação possa auxiliar a endereçar os desafios debatidos acima, é preciso ir além das regras de privacidade e se construir uma verdadeira agenda através dos princípios éticos da inteligência artificial e das decisões automatizadas amparadas no uso de algoritmos, [...].¹¹⁴

¹¹³ Sobre *big data*: O termo *big data* refere-se às possibilidades de acesso a grandes quantidades de dados de diferentes tipos, qualidades e formas de coleta ("volume"), bem como alta velocidade de processamento ("velocity"). Além disso, o *big data* é a base de novos modelos de negócios e possibilidades de várias criações de valor ("value"), na medida em que pode ser usado em conjunto com outras tecnologias como a Internet das Coisas ou *Cloud Computing*. *Ibidem*, p. 5.

¹¹⁴ DONEDA, Danilo Cesar Maganhoto *et al.* Considerações iniciais sobre inteligência artificial, ética e autonomia pessoal. **Pensar – Revista de Ciências Jurídicas**, Fortaleza, v. 23, n. 4, p. 1-17, out/dez. 2018. DOI: 10.5020/2317-2150.2018.8257. Disponível em: <https://ojs.unifor.br/rpen/article/view/8257>. Acesso em: 05 out. 2022. p. 6.

Em semelhante sentido, O'Neil afirma que a substituição de análise humana por aplicações matemáticas acaba sendo baseada em escolhas humanas falíveis, mesmo quando não são intencionais. Dentre os principais problemas indicados pela autora, um dos motivos destas ocorrências também é a utilização de *proxies*¹¹⁵ quando não há dados mensuráveis para tomada de determinada decisão.

A gravidade de tal circunstância é alertada por O'Neil:

[...]. Ao longo do caminho, testemunhamos a destruição causada pelas ADMs. Prometendo eficiência e justiça, elas distorcem o ensino superior, aumentam as dívidas, estimulam o encarceramento em massa, esmagam os pobres em quase todos os momentos, e minam a democracia. Pode parecer que a resposta lógica seria desarmar essas armas, uma por uma.

O problema é que elas alimentam-se umas das outras. Pessoas pobres são mais propensas a ter pouco crédito e viver em bairros com maior incidência de crimes, cercadas por outras pessoas pobres. Uma vez que o universo sombrio das ADMs digere esses dados, ele inunda essas pessoas com anúncios predatórios de empréstimos imobiliários de alto risco ou universidades com fins lucrativos. Envia mais policiamento para prendê-las, e quando são condenadas as sentença com penas mais duras. Esses dados alimentam outras ADMs, que marcam as mesmas pessoas como de alto risco ou alvos fáceis e prosseguem a bloqueá-las empregos, enquanto aumentam seus juros para empréstimos imobiliários, de carros e todo e qualquer plano de seguro imaginável. Isso derruba ainda mais sua classificação de crédito, criando nada menos que uma espiral moral de modelagem. Ser pobre em um mundo de ADMs está se tornando cada vez mais perigoso e custoso.¹¹⁶

¹¹⁵ O conceito de *proxy* já fora abordado anteriormente neste trabalho, o qual pode ser complementado com a visão de Cathy O'Neil, que esclarece que em razão da carência de dados para os comportamentos em que se tem interesse, os programadores acabam utilizando dados substitutos ou indicadores aproximados (*proxies*). Ela afirma que eles traçam *correlações* estatísticas entre o CEP residencial de alguém ou padrões de linguagem e seu potencial de pagar um empréstimo ou conseguir dar conta de um emprego. Segundo Cathy, estas correlações são discriminatórias e, algumas delas, ilegais. O'NEIL, Cathy. **Algoritmos de destruição em massa**: como o big data aumenta a desigualdade e ameaça à democracia. Tradução Rafael Abraham. 1. ed. Santo André/SP: Rua do Sabão, 2020. p. 20.

¹¹⁶ Cabe um breve esclarecimento: a sigla ADM significa Armas de Destruição em Massa, que é como a autora chama os algoritmos discriminatórios. O'NEIL, Cathy. **Algoritmos de destruição em massa**: como o big data aumenta a desigualdade e ameaça à democracia. Tradução Rafael Abraham. 1. ed. Santo André/SP: Rua do Sabão, 2020. p. 223.

Esta “espiral moral de modelagem” é também chamada pela autora de “ciclo de *feedback*” ou “retroalimentação”, pois um julgamento errado da IA acaba gerando novos dados que também estão errados e que vão embasar novos julgamentos que se refletirão em estatísticas novamente erradas. E o ciclo em *looping* se torna infinito, cristalizando aquela realidade.

Imagine hipoteticamente se determinada empresa resolvesse treinar carros autônomos a partir de dados de todos os acidentes de trânsito ocorridos. Se o percentual maior de óbitos ocorresse com vítimas crianças, o carro inteligente poderia “aprender” que, em caso de acidente envolvendo uma criança e um idoso, ele deve desviar e proteger o idoso mesmo que corra o risco de atingir e matar a criança, já que, em regra, ela pertence a um grupo de pessoas que falece com mais frequência neste tipo de acidente. E o pior é que esta nova morte entrará para estatística de crianças acidentadas e reforçará o “aprendizado” da máquina.

Além das causas acima relatadas, destaca-se que o desenvolvimento dos principais *softwares* e algoritmos está no domínio de poucas *big techs*¹¹⁷, as maiores empresas de tecnologia do mundo, muitas delas situadas no Vale do Silício, comandadas majoritariamente por homens brancos, cujas características sociais privilegiadas podem cristalizar alguns poucos vieses que dominarão toda sociedade. Sobre o ponto, transcreve-se a reflexão:

¹¹⁷ Juarez Freitas e Thomas Freitas atentam, inclusive, para a necessidade de regulação deste domínio mercadológico: “No tocante ao abuso do poder econômico e à usurpatória posição dominante de mercado, a regulação da IA precisa diluir a concentração desmesurada de poder, provocada pela IA, em escala inédita. Não se trata de combater os monopólios naturais, mas de desbaratar as redes de dominância sobrepostas aos Estados, num ambiente em que as fronteiras esmaecem. Nesse quadro de acentuada desmaterialização, justifica-se a regulação que protege, vigorosamente, a liberdade contra qualquer subjugação virtual, por ação ou omissão, nacional ou transnacional.” FREITAS, Juarez; FREITAS, Thomas Bellini. **Direito e Inteligência**

Artificial: em defesa do humano. 1. ed. Belo Horizonte: Forum, 2020. p. 59. Bruno Bioni e Maria Luciano também denunciam: “Dentre as razões para isso têm sido apontadas a falta de regulação, monopólios no setor de IA, estruturas de governança insuficientes dentro de empresas de tecnologia, assimetrias de poder entre empresas e usuários, a distância cultural entre os responsáveis por pesquisas em tecnologia e a diversidade das populações nas quais essa tecnologia é utilizada (AI NOW, 2018). Esse diagnóstico tem suscitado demandas sociais por maior transparência no uso de IA.” BIONI, Bruno Ricardo; LUCIANO, Maria. O princípio da

precaução na regulação de inteligência artificial: seriam as leis de proteção de dados o seu portal de entrada. In: MULHOLLAND, Caitlin; FRAZÃO, Ana (Coord.). **Inteligência artificial e direito:** ética, regulação e responsabilidade. São Paulo: Thomson Reuters Brasil, 2019. *E-book*.

Nota-se, assim, que, na própria constituição dos sistemas de IA se fazem escolhas que refletem também as opiniões e prioridades dos criadores, as quais influenciam diretamente as respostas do sistema. Não se pode ignorar, assim, a impossibilidade de isenção completa, até mesmo ao se falar de inteligência artificial e de sistemas que, muitas vezes, são tratados como universais e “desenviesados”, porquanto o ponto de partida é sempre uma atividade humana de seleção de informações e dados, os quais refletem, também, o contexto social de quem os produziu. E tal preocupação se liga diretamente ao estudo dos vieses cognitivos (cognitive biases).¹¹⁸

E tudo isso ocorre sem as pessoas jamais saberem o que embasou determinado feito sobre si, até porque normalmente nem se sabe que se está sendo julgado por um algoritmo.

Em suma, pode-se agrupar basicamente em dois grandes grupos as principais causas da discriminação algorítmica:

a) **Discriminação por dados:** considerando que as máquinas aprendem com base nas informações/dados que recebem para analisar e que as informações atualmente produzidas pela humanidade são tomadas de preconceitos e refletem discriminações, a máquina toma aquilo como premissa e replica o que “vê”. Tal situação é ilustrada pela famosa frase “lixo entra, lixo sai”.¹¹⁹ A discriminação também pode ocorrer quando os dados fornecidos ao sistema são insuficientes ou não refletem a realidade, desconsiderando grupos ou fatos;

b) **Discriminação por escolhas humanas:** o conhecimento técnico para programar algoritmos está limitado a um grupo dotado de pouca diversidade, uma vez que é ocupado em sua maioria por homens brancos¹²⁰ que, ainda que de boa-fé,

¹¹⁸ NUNES, Dierle; MARQUES, Ana Luiza. Inteligência artificial e direito processual: vieses algorítmicos e os riscos de atribuição de função decisória às máquinas. *Revista de Processo*, [S.l.], v. 285, p. 421-447, nov. 2018.

¹¹⁹ A expressão GIGO (*Garbage in, garbage out*) é atribuída ao técnico da IBM George Fuechsel.

¹²⁰ A informação tem como base a pesquisa realizada em 2021 pelo fórum StackOverFlow com mais de 80.000 desenvolvedores: FÓRUM Stack Overflow. **Developer survey**. 2021. Disponível em: <https://insights.stackoverflow.com/survey/2021#developer-profile-demographics>. Acesso em: 15 set. 2022, e pela ARTIFICIAL INTELLIGENCE INDEX – AI; STANFORD INSTITUTE FOR HUMAN-CENTERED ARTIFICIAL INTELLIGENCE. **The AI Index Report, 2021**. Disponível em: <https://aiindex.stanford.edu/report/#individual-chapters>. Disponível em: 10 out. 2022.

não possuem experiências de vida que lhes permitam enxergar o mundo por todos os ângulos necessários, fazendo com que escolham a seu critério os *proxys* e/ou as generalizações que indicarão os resultados a serem perseguidos pelos algoritmos, mesmo que sem intencionalidade, na medida em que refletem vieses históricos. As escolhas humanas também podem provocar discriminação quando rotularem de forma equivocada ou enviesada os dados e variáveis que orientarão as regras sistêmicas;

Ainda, é de se considerar que denúncias sobre a ocorrência de vieses discriminatórios perpetrados por inteligência artificial vêm sendo apontadas há anos, sendo publicadas a partir de pesquisas acadêmicas, investigações jornalísticas, já chegando ao público em geral até mesmo por meio de documentários disponibilizados no Netflix (*Coded Bias*, *Privacidade Hackeada*, *o Dilema das Redes* etc.) e vídeos no Youtube (há diversos *TED Talks* abordando esta temática).

Portanto, não se pode alegar desconhecimento dos riscos de se utilizar inteligência artificial sem observância de princípios éticos, o que exige das empresas que se utilizam desses sistemas que ajam com a devida diligência antes, durante e depois da tomada de decisões com base nestas aplicações.

Dada a ausência de regulamentação vigente que discipline o desenvolvimento e o uso da inteligência artificial no Brasil e em diversos países do mundo, em virtude da dificuldade de compreender seu funcionamento, bem como da ausência de consenso com relação aos princípios éticos que precisam ser observados em todo o ciclo de vida desta tecnologia, o regramento está delegado à discricionariedade das empresas que dispõem destas aplicações, submetendo seu uso a toda população mundial sem qualquer tipo de obrigação legal para obstar seus efeitos indesejados.

Neste sentido, tendo em vista que tais riscos podem ensejar a violação de direitos fundamentais e até mesmo a prática de crimes, é cogente que se avalie se as corporações que desenvolvem ou se utilizam de inteligência artificial devem implementar políticas internas organizando todo o ciclo de vida da IA, se valendo dos princípios de governança durante todas as etapas do uso destas aplicações, mesmo que não haja disposição legal direta impondo esta obrigatoriedade, o que se fará a seguir.

É importante destacar que, consoante uma pesquisa realizada pela Amy Webb, o desenvolvimento de inteligência artificial está basicamente na mão de nove grandes companhias, as quais chama de "*The Big Nine*", sendo elas: Google, Amazon, Apple, IBM, Microsoft e Facebook nos Estados Unidos e Baidu, Alibaba e Tencent na China.¹²¹ Em semelhante sentido, Hoffmang-Riem também refere que "o desenvolvimento da IA é dominado em grau crescente por grandes empresas de TI e empresas especializadas a elas associadas."¹²² Harari também atenta para a concentração de poder e dados em empresas como Google e Facebook.¹²³

Por este motivo, entendeu-se relevante investigar na legislação empresarial se há algum dever de adotar cautelas mínimas no uso de inteligência artificial por empresas como estas, no intuito de mitigar os riscos que podem se originar a partir de tais aplicações, o que será avaliado a seguir, com foco no dever de diligência que os administradores das sociedades anônimas devem adotar na condução de seus negócios.

2.9 Conclusões parciais

Consolidando e resumindo o que foi pesquisado sobre inteligência artificial envolvendo aprendizagem de máquina, pôde-se constatar que a inteligência artificial está em constante expansão nas rotinas empresariais e pessoais, sobretudo em virtude da possibilidade de (i) identificar padrões, (ii) correlacioná-los e (iii) classificar grupos de características comuns, aplicando tais conclusões de forma estatística em eventos futuros, permitindo que os sistemas "adivinhem" situações e realizem tarefas com muita facilidade.

Todavia, tais benefícios surgem acompanhados de determinados efeitos colaterais, que podem influenciar negativamente a vida de pessoas e grupos, de

¹²¹ WEBB, Amy. **The Big Nine**: how the tech titans and their thinking machines could warp humanity. New York: PublicAffairs, 2019. p. 11.

¹²² HOFFMANN-RIEM, Wolfgang. Inteligência artificial como oportunidade de regulação jurídica. **Revista de Direito Público**, Brasília, n. 90, v. 16, p. 11-38, nov./dez. 2019.

¹²³ HARARI, Yuval Noah. **Homo Deus**: uma breve história do amanhã. São Paulo: Companhia das Letras, 2016. p. 394.

forma que isso restrinja direitos e garantias individuais, como o direito à igualdade, configurando o que se convém chamar de discriminação algorítmica.

A discriminação algorítmica, por sua vez, foi categorizada em grupos por diversos autores analisados, a partir de diferentes premissas, tais como suas origens ou seus efeitos. Tais categorizações, apesar de não serem unânimes, evidenciaram que cada etapa do desenvolvimento de um sistema de inteligência artificial, especialmente os que utilizam aprendizagem de máquina, deve exigir um determinado tipo de atenção e cuidados para que não seja prejudicial e discriminatória a terceiros.

Tal constatação exige dos agentes que utilizam e/ou ofertam inteligência artificial a criação de mecanismos que garantam a adoção de cautelas específicas, na tentativa de evitar, ou ao menos mitigar, a perpetração de vieses discriminatórios negativos. Tais cautelas podem exigir pesquisa, investimento de tempo e dinheiro, envolvimento de partes interessadas, reconhecimento (e até exposição) de vulnerabilidades, contratação de terceiros para assessoramento e auditorias e/ou até prejuízos e retrabalho.

Neste sentido, a decisão de implementar e como implementar inteligência artificial nas rotinas empresariais não deve estar somente com o setor de tecnologia da informação, mas ser encarada como uma decisão estratégica que pode influenciar diretamente nos negócios, cujas orientações e disseminação da cultura de cuidados partam da própria gestão.

Até mesmo porque, na medida em que se delegam às máquinas atividades e decisões que seriam tomadas por trabalhadores, a empresa se sujeita a correr todos os riscos legais, financeiros e reputacionais que podem decorrer de algum resultado discriminatório. E nenhum administrador pode correr este risco sem antes empregar máximos cuidados e diligências, como exige a previsão do art. 153 da Lei 6.404/76.

É neste sentido que, passo seguinte, avaliar-se-á quais são as condutas que representam o cuidado necessário para que seja reconhecido o cumprimento do dever de diligência, no âmbito das empresas que utilizam inteligência artificial.

3. DO DEVER DE DILIGÊNCIA

3.1 Exame da doutrina

O objetivo deste capítulo é investigar a interpretação dada pela doutrina ao dever de diligência, previsto no artigo 153 da Lei de Sociedades Anônimas (Lei 6.404/1976, doravante denominada Lei das SA), o qual estipula que “O administrador da companhia deve empregar, no exercício de suas funções, o cuidado e diligência que todo homem ativo e probo costuma empregar na administração dos seus próprios negócios.”¹ A mesma redação é repetida no artigo 1.011 do Código Civil Brasileiro, aplicável às sociedades simples e limitadas.²

O apontamento acima se faz necessário, pois, segundo pesquisas, grande parte das empresas de tecnologia sediadas no Brasil encontram-se sobre a forma de sociedade limitada, inclusive algumas das *big techs*, e é sobre estas empresas que recai o tema desta pesquisa.³

Dessa forma, as interpretações sobre esse dispositivo também podem ser plenamente aproveitadas para outras configurações societárias⁴, como o caso das sociedades limitadas, em razão da identidade na previsão legal do dever de

¹ BRASIL. **Lei n. 6.404, de 15 de dezembro de 1976**. Dispõe sobre as sociedades por ações. Brasília: Casa Civil, 1976. Disponível em: https://www.planalto.gov.br/ccivil_03/Leis/L6404consol.htm. Acesso em: 15 abr. 2022.

² Código Civil: Art. 1.011. O administrador da sociedade deverá ter, no exercício de suas funções, o cuidado e a diligência que todo homem ativo e probo costuma empregar na administração de seus próprios negócios. BRASIL. **Lei 10.406, de 10 de janeiro de 2002**. Institui o Código Civil. Brasília: Presidência da República, 2002. Disponível em: https://www.planalto.gov.br/ccivil_03/LEIS/2002/L10406compilada.htm. Acesso em: 10 out. 2022.

³ DIAS, Natasha. Panorama da adoção de tecnologia por empresas brasileiras. **Cortex**, 30 ago. 2022. Disponível em: <https://www.cortex-intelligence.com/intelligence-review/panorama-sobre-adoacao-de-tecnologia-por-empresas-brasileiras>. Acesso em: 17 out. 2022; SOUZA, Ludmilla. Indústria de software e serviços de TIC cresceu 6,5% em 2021. **Agência Brasil**. 19 jul. 2022. Disponível em: <https://agenciabrasil.ebc.com.br/economia/noticia/2022-07/industria-de-software-e-servicos-de-tic-cresceram-65-em-2021>. Acesso em: 17 out. 2022; ASSOCIAÇÃO PARA PROMOÇÃO DA EXCELÊNCIA DO SOFTWARE BRASILEIRO – SOFTEX. **Observatório Softex 2019**. Overview do setor de tecnologia da informação brasileiro nos últimos dez anos. Brasília, 2019. *E-book*. Disponível em: <https://softex.br/download/overview-do-setor-de-tecnologia-da-informacao-brasileiro-nos-ultimos-dez-anos/?wpdmdl=97749&masterkey=5cee9abb82a39>. Acesso em: 17 out. 2022.

⁴ FRAZÃO, Ana. Responsabilidade civil dos administradores de sociedades empresárias por decisões tomadas com base em sistemas de inteligência artificial. In: FRAZÃO, Ana; MULHOLLAND, Caitlin (Coord.). **Inteligência artificial e Direito: ética, regulação e responsabilidade**. 2. ed. rev. atual. e ampl. São Paulo: Thomson Reuters Brasil, 2020. p. 520. *E-book*.

diligência. Todavia, a tradição da doutrina despeja mais foco na análise do artigo 153 da Lei das SA, motivo pelo qual a pesquisa se concentrará nele.

O citado dispositivo da Lei das SA impôs um dever aos administradores, sem, contudo, indicar de forma prática e objetiva quais são as condutas que considera como sendo diligentes, deixando a interpretação de seu cumprimento à grande subjetividade e discricionariedade.

Em razão da imprecisão dos conceitos eleitos pelo legislador, tais como 'cuidado', 'diligência', 'ativo' e 'probo', abre-se espaço para atribuir a seu conteúdo as mais diversas definições, que certamente se modificaram ao longo do tempo. Isso porque as tendências corporativas de cada época e as ferramentas científicas e tecnológicas que propiciam a busca por dados, informações e investigações podem alterar a percepção do que é ser diligente.

As condutas esperadas de um administrador diligente do ano de 1976 certamente não são as mesmas que se esperam de um administrador em 2023. Nestes quase cinquenta anos, surgiram novas formações, cursos, especialidades, ferramentas, internet, tecnologias, diferentes formas de se comunicar etc., que modificaram (e muito) o mercado profissional em todas as áreas.

Deste modo, para esta pesquisa, busca-se compreender (i) se o teor do art. 153 da Lei das SA é mesmo adaptável conforme as necessidades do contexto em que a empresa e o administrador estão inseridos, bem como (ii) quais são as características e ações que um administrador deve possuir ou adotar ao gerir uma empresa que se utiliza de inteligência artificial em suas rotinas, para que seja considerado diligente.

Neste sentido, diversos autores e autoras se dedicaram a identificar o que significa ser um administrador diligente cumpridor deste dever e que tipo de ação ou omissão pode ensejar o seu descumprimento. Os autores estudados para este capítulo foram: Flavia Parente, Renato Ventura Ribeiro, Daniele de Lima Oliveira, Luiz Antonio de Sampaio Campos, Otavio Yazbek, Pedro Henrique Castello Brigagão, Larissa Quattrini, Nelson Eizirik, Alexandre Couto e Silva, Maria Cristina Borrasca Felisberto, Carlos Konrad, Ricardo Lupion, Laís Machado Lucas, Rodrigo Castro, Ana Frazão, Maria Eduarda Ferreira Piccoli e Oksandro Osdival Gonçalves, cujas

conclusões serão discorridas adiante em ordem cronológica, no intuito de acompanhar a evolução da interpretação deste dever.

Flavia Parente acena que o dever de diligência é de difícil definição em razão da opção de diversos legisladores adotarem um conceito em abstrato, recorrendo a padrões de comportamento (a exemplo do adotado pela legislação brasileira: homem ativo e probo), ressaltando, no entanto, que a própria palavra diligência já traz a ideia de zelo, esforço e cuidado no cumprimento de um dever.⁵

Refere a autora que o dever de diligência tem caráter instrumental ou acessório.⁶ Inicialmente, as expectativas com relação a tal dever ficavam atreladas ao comportamento de um *bonus pater familias* (bom pai de família), que desempenharia suas funções com atenção, zelo e cuidado. Tendo em vista se tratar de um paradigma genérico e subjetivo, cuja interpretação pode variar conforme o contexto analisado, momento, local etc., a análise do cumprimento do dever de diligência dependerá de cada caso concreto.⁷

Todavia, com o passar dos anos e com a evolução e profissionalização das empresas, a expectativa sobre o que significa "cumprir o dever de diligência" se

⁵ PARENTE, Flávia. **O dever de diligência dos administradores de sociedades anônimas**. Rio de Janeiro: Renovar, 2005. p. 37.

⁶ PARENTE, Flávia. **O dever de diligência dos administradores de sociedades anônimas**. Rio de Janeiro: Renovar, 2005. p. 41.

⁷ No mesmo sentido: Daniele de Lima Oliveira examinou o dever de diligência e entendeu que o art. 153 da Lei 6404/76 prescreve padrões de comportamento amplos e flexíveis, de modo a englobar diversas situações. Na visão da autora, a boa-fé é o princípio norteador dos parâmetros estipulados. Admite que, apesar de a lei brasileira ter adotado um preceito vago, seguiu o exemplo de diversas outras legislações, todas inspiradas no conceito do "*bonus pater familias*", que atrai a ideia de cuidado e zelo. Ainda refere que, mais recentemente, legislações passaram a equiparar este parâmetro com o do homem de negócios ativo e probo. Neste sentido, interpreta a autora que ser diligente é empregar todas "as cautelas, os métodos, recomendações, postulados e diretivas da "ciência" da administração de empresas", destacando que seu cumprimento é uma obrigação de meio e não de resultado, a qual exige do "administrador uma postura proativa na fiscalização da companhia, inclusive no que se refere à fiscalização da observância dos princípios da governança corporativa". OLIVEIRA, Daniele de Lima. **Deveres e responsabilidade dos administradores da S/A**. Dissertação (Mestrado em Direito) - Faculdade de Direito da Pontifícia Universidade Católica de São Paulo, São Paulo, 2008. p. 56. No mesmo sentido: Pedro Henrique Castello Brigagão constata que o art. 153 da Lei 6.404/72 estabelece um conceito amplo e genérico, elegendo a figura do *bonus pater familias*, oriunda do Direito Romano, como modelo de conduta. Comenta acerca das críticas doutrinárias a respeito da opção do legislador, pelo fato de o pai de família ter características mais conservadoras, objetivando a preservação da entidade familiar, a prudência, enquanto o empresário objetiva a obtenção de lucros, de expansão do negócio, exigindo cerca profissionalização e especialização. BRIGAGÃO, Pedro Henrique Castello. **Dever de diligência dos administradores de sociedades anônimas**. Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Direito. Rio de Janeiro, 2012. Disponível em: https://www.puc-rio.br/ensinopesq/ccpg/pibic/relatorio_resumo2012/resumos_pdf/ccs/DIR/Pedro%20Henrique%20Castello%20Brigag%20E3o.pdf. Acesso em: 17 out. 2022.

ajustou à realidade atual, exigindo-se atuação competente e profissional dos administradores, pois a cautela que se exige de um “bom pai de família” não estava compatível com a moderna administração societária.⁸

Deste modo, segundo Parente, o cumprimento do dever de diligência necessita de análise das situações concretas, avaliando se determinados comportamentos foram adotados pelo administrador antes ou durante determinada ação ou omissão, ainda que o objetivo final da companhia não tenha sido plenamente alcançado.⁹

A pesquisadora demonstra que parte da doutrina, como Fabio Ulhôa Coelho, Modesto Carvalhosa e Fran Martins, passou a defender a exigência de competências profissionais específicas, adquiridas por escolaridade e/ou experiência, investindo o administrador de conhecimentos mínimos que o tornem diligente. Contudo, destaca que, mesmo possuindo tais capacidades, incumbiria verificar “caso a caso, qual seria a atitude recomendável, naquelas circunstâncias específicas, naquele tipo de negócio, de acordo com as normas da ciência da administração de empresas.”¹⁰

Segundo Parente, se, ao avaliar o contexto das ações tomadas pelo administrador, constatar-se que ele agiu de modo adequado perante a situação que se encontrava, significa que foi cumprido o dever de diligência, não podendo responsabilizar o administrador em caso de eventual resultado prejudicial.¹¹ A autora concluiu que o dever de diligência se decompõe em cinco deveres distintos, sendo

⁸ Luiz Antonio Sampaio Campos destaca que há diferenças nos padrões de conduta adotados por diferentes legislações, o que muitas vezes suscita discussões. Aduz que tal diferença não produz grandes divergências na hora de interpretar e julgar o (des)cumprimento do dever de diligência. Porém, indica que, diferente do que muitos alegam, o legislador brasileiro não adotou o comportamento do *pater familiae* como base, mas, sim, do homem de negócios, por entender ser este mais apropriado do que o homem comum. Comenta ser um padrão mais rigoroso, eis que, enquanto o bom pai de família remete a uma pessoa conservadora e avessa a riscos, o administrador precisa ter aptidão ao risco empresarial e ousadia para adotar posturas criativas e inovadoras. CAMPOS, Luiz Antonio de Sampaio. Deveres e responsabilidades. In: LAMY FILHO, Alfredo; PEDREIRA, José Luiz Bulhões (Coord.). **Direito das companhias**. São Paulo: Forense, 2009. Rodrigo Castro também entende que o parâmetro escolhido pela Lei das SA foi o(a) empresário(a) probo(a), definindo, “isto é, íntegro e honesto, mas, por definição, capaz de conviver, com alguma naturalidade, em ambientes competitivos e estressados e, de modo informado, alocar riscos.”. Aduz que a Comissão de Valores Mobiliários já afastou a ideia de que o dever de diligência refere o padrão de conduta do *bonus pater familiae*, por entender que o padrão de conduta eleito pela Lei das SA é de um profissional com capacidades técnicas. COELHO, Fábio Ulhoa (Coord.) et al. **Lei das sociedades anônimas comentada**. Rio de Janeiro: Forense, 2021. p. 873 e segs.

⁹ PARENTE, Flávia. **O dever de diligência dos administradores de sociedades anônimas**. Rio de Janeiro: Renovar, 2005. p. 50.

¹⁰ *Ibidem*, p. 52.

¹¹ *Ibidem*, p. 52.

eles os deveres de: se qualificar, bem administrar, se informar, investigar e vigiar, os quais serão explicados adiante.

Já Renato Ventura Ribeiro conceitua o dever de diligência como:

O termo diligência provém do latim *diligere*, que significa zelar ou cuidar. No âmbito jurídico, diligência é o adequado emprego de energias e meios úteis para realização de um fim determinado. Portanto, diligência traz o modo de conduta a ser observado. Trata-se de uma obrigação legal e contratual.

No sentido subjetivo, o dever de diligência pode ser tido como esforço, dedicação, cuidado, interesse, atenção e zelo na execução de uma tarefa, independentemente do resultado final. E numa concepção objetiva, o padrão de conduta exigido no exercício de uma atividade.¹²

Em seu livro, denota-se que o dever de diligência é uma obrigação de meio¹³, pois o autor entende que o “administrador deve agir com a diligência que melhor sirva para obter o melhor resultado possível, [...]”¹⁴ e que, “superando obrigações meramente formais, é um dever de prudência no exercício do cargo e de atuação conforme critérios geralmente admitidos em cada setor de atividade.”¹⁵

Conforme Ribeiro, para apurar se um administrador honrou com o dever de diligência, deve-se averiguar dois critérios de apuração - subjetivo e objetivo:

¹² RIBEIRO, Renato Ventura. **Dever de diligência dos administradores de sociedades**. São Paulo: Quartier Latin, 2006. p. 227-228.

¹³ Na mesma linha, Larissa Quattrini, Nelson Eizirk e Ana Frazão também reconhecem que o dever de diligência se trata de uma obrigação de meio e não de resultado. QUATTRINI, Larissa Teixeira. **Os deveres dos administradores de sociedades anônimas abertas: estudo de casos**. São Paulo: Saraiva, 2014. *E-book*.; EIZIRIK, Nelson. A lei das S/A comentada. 2. ed. rev. e ampl. Artigos 138-205. São Paulo: Quartier Latin, 2015. V. II, p. 115-124; COELHO, Fábio Ulhoa (Coord.) et al. **Lei das sociedades anônimas comentada**. Rio de Janeiro: Forense, 2021. p. 115-124. Com foco no uso de inteligência artificial, Maria Cristina Borrasca Felisberto investiga o limite dos riscos assumidos pelos administradores, aduzindo que o art. 153 da Lei das SA, ao não estabelecer um rol exaustivo de condutas, exigiu do administrador uma diligência ordinária, que deve ser avaliada a cada caso, considerando-se as atividades da companhia, seu porte, o contexto em que a decisão foi tomada, quais as informações buscadas pelo administrador, comparando-o com outros em situação semelhante. Reforça que a legislação, ao impor um dever de conduta, exige do administrador uma obrigação de meio e não de resultado, classificando uma decisão como diligente ou não a partir do comportamento adotado durante o processo de convencimento.

¹⁴ RIBEIRO, Renato Ventura. **Dever de diligência dos administradores de sociedades**. São Paulo: Quartier Latin, 2006. p. 210.

¹⁵ *Ibidem*, p. 215.

Pelo critério subjetivo apura-se a culpabilidade da conduta do agente, se dolosa ou apenas culposa, bem como se ele tinha eventual previsão das consequências dos atos práticos. Mas a culpa não deve incluir a culpa levíssima, na qual somente não incorreria um homem excepcionalmente diligente, salvo se o administrador foi contratado por tais características e correspondente remuneração.

Através do critério objetivo verifica-se a observância das leis, dos estatutos e contratos sociais, de boa-fé, dos demais contratos e dos usos e costumes empresariais, entre outros.¹⁶

Luiz Antonio de Sampaio Campos leciona acerca da opção do legislador brasileiro de trazer o rol de deveres esperados dos administradores das sociedades anônimas, bem como das responsabilidades também a eles imputadas. O autor refere que os deveres surgem antes das responsabilidades e demonstra que a lei brasileira se deu de modo diverso das legislações de outros países – que elencaram somente as hipóteses de responsabilização e delegaram ao intérprete a tarefa de compreender os deveres.¹⁷

Com relação ao dever de diligência, Campos afirma que este dever é o de maior abrangência imposto pela Lei das SA, sendo a “pedra de toque da atuação dos administradores.” Por isso, o dever configura, segundo o autor, “a ‘chave’ da responsabilização dos administradores.”¹⁸

Campos defende que a escolha do padrão de conduta do homem ativo e probo traz importante flexibilidade para a interpretação do que se espera do dever de diligência, permitindo que ele se ajuste às realidades das sociedades anônimas. Afirma que “adotou-se, então, um conceito indeterminado, mediante um modelo típico e abstrato – e não de um caso particular – que permitirá ao intérprete dar conteúdo concreto ao dever de diligência, conforme as circunstâncias e de acordo com a estrutura da companhia.”¹⁹

¹⁶ *Ibidem*, p. 218-219.

¹⁷ CAMPOS, Luiz Antonio de Sampaio. Deveres e responsabilidades. *In*: LAMY FILHO, Alfredo; PEDREIRA, José Luiz Bulhões (Coord.). **Direito das companhias**. São Paulo: Forense, 2009. p. 1085.

¹⁸ CAMPOS, Luiz Antonio de Sampaio. Deveres e responsabilidades. *In*: LAMY FILHO, Alfredo; PEDREIRA, José Luiz Bulhões (Coord.). **Direito das companhias**. São Paulo: Forense, 2009. p. 1097.

¹⁹ *Ibidem*, p. 1099.

Em sua pesquisa, o autor conclui que a análise do cumprimento do dever de diligência deve passar por uma análise subjetiva, avaliando a capacidade e a perícia de cada administrador, para que se constate como ele deveria/poderia ter agido:

A diligência que se exige do administrador é a diligência ordinária; não se requer uma diligência excepcional. Da mesma forma, a diligência não é uniforme ou geral. Será um padrão de diligência caso a caso, levando-se em consideração as características da companhia, da decisão, do administrador, da informação disponível, do tempo, dos recursos disponíveis etc.

Por isso é que se diz que a diligência não é só um critério abstrato e genérico, mas também relativo e adaptável às circunstâncias. Sua aplicação deve ser concreta, à vista da realidade do caso.

[...]

Não existe, entretanto, uma única forma de atuar diligentemente, mas uma zona de comportamentos que podem ser considerados diligentes, o que significa dizer, por outras palavras, que, para fins de verificar o cumprimento desse dever por parte dos administradores, não se deve indagar qual a conduta mais diligente, mas, sim, se houve falta de diligência no comportamento adotado. Naturalmente, a existência de uma conduta mais recomendável não significa que uma outra conduta também não possa ser considerada diligente, notadamente quando se tem presente o conceito do tipo médio.²⁰

Em semelhante sentido, Otavio Yazbek refere que o art. 153 da Lei das SA é uma cláusula geral, estrategicamente utilizada para permitir variações de forma e de conteúdo.²¹ Também reconhece as vantagens de tal modelo, como por exemplo evitar o risco de se prescrever comportamentos muito específicos, que se tornem inadequados com o passar do tempo, tendo em vista que as boas práticas de gestão podem mudar conforme o momento e as características de cada empresa. Contudo, adverte que a elasticidade do modelo utilizado pode trazer inseguranças quanto à interpretação que será dada na análise do caso concreto.

²⁰ *Ibidem*, p. 1101-1102.

²¹ YAZBEK, Otavio. Representação do dever de diligência na doutrina jurídica brasileira: um exercício e alguns desafios. In: KUYVEN, Luiz Fernando Martins (Coord.). **Temas essenciais de direito empresarial**: estudos em homenagem a Modesto Carvalhosa, São Paulo: Saraiva, 2012. E-book.

Com a intenção de delinear concretamente o que consiste no cumprimento do dever de diligência, Pedro Henrique Castello Brigagão examina legislação, doutrina especializada e precedentes norte-americanos, bem como da Comissão de Valores Mobiliários (CVM), sobre o tema.²² A partir disso, Brigagão descreve os diversos aspectos que, para ele, compõem o dever de diligência, sendo eles: o dever de se informar, o dever de se qualificar, o dever de vigilância ou dever de investigar, o dever de intervir, bem como os deveres menos citados pela doutrina, como os de bem administrar, de não praticar erros graves e de assiduidade/participação.

O pesquisador elaborou três conclusões: os administradores devem (i) ter precaução com relação às informações que recebem de terceiros, mesmo quando estes são outros administradores, (ii) ter prudência, mesmo quando se tratar de situações que demandem urgência, estudando e analisando as variáveis antes de tomar decisões e (iii) se basear em estudos de especialistas. Ao final de sua pesquisa, Brigagão afirma que o dever de diligência é violado quando há uma conduta negligente, imprudente ou imperita. Ou seja, o autor vincula o descumprimento do dever de diligência com condutas culposas, quando lhes falta cuidado, atenção e zelo com as necessidades e atividades da companhia.

Larissa Quattrini optou por investigar os deveres dos administradores das sociedades anônimas a partir do estudo de casos concretos. A autora defende que outros dispositivos da lei societária, tal como o art. 154 da Lei da SA²³, estabelecem

²² BRIGAGÃO, Pedro Henrique Castello. **Dever de diligência dos administradores de sociedades anônimas**. Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Direito. Rio de Janeiro, 2012. Disponível em: https://www.puc-rio.br/ensinopesq/ccpg/pibic/relatorio_resupdf. Acesso em: 17 out. 2022.

²³ Art. 154. O administrador deve exercer as atribuições que a lei e o estatuto lhe conferem para lograr os fins e no interesse da companhia, satisfeitas as exigências do bem público e da função social da empresa.

§ 1º O administrador eleito por grupo ou classe de acionistas tem, para com a companhia, os mesmos deveres que os demais, não podendo, ainda que para defesa do interesse dos que o elegeram, faltar a esses deveres.

§ 2º É vedado ao administrador:

a) praticar ato de liberalidade à custa da companhia;
b) sem prévia autorização da assembléia-geral ou do conselho de administração, tomar por empréstimo recursos ou bens da companhia, ou usar, em proveito próprio, de sociedade em que tenha interesse, ou de terceiros, os seus bens, serviços ou crédito;
c) receber de terceiros, sem autorização estatutária ou da assembléia-geral, qualquer modalidade de vantagem pessoal, direta ou indireta, em razão do exercício de seu cargo.

§ 3º As importâncias recebidas com infração ao disposto na alínea c do § 2º pertencerão à companhia.

os padrões de conduta esperados, os quais, segundo ela, demonstram a “predileção do legislador brasileiro pelo princípio da boa-fé”²⁴ e almejam atingir os fins previstos no art. 154.²⁵

Nelson Eizirik identifica em decisões e orientações norte-americanas os deveres de participar, de estar bem-informado (resguardando-se o direito de confiar nas informações que recebe quando não há motivos para o contrário), de investigar e de informar seus pares. Após, complementa apontando os deveres de se qualificar para o cargo, de bem administrar e vigiar.²⁶

Alexandre Couto Silva afirma que “entre os principais princípios da governança corporativa estão os deveres fiduciários dos administradores de sempre agirem de maneira informada, de boa-fé com o devido cuidado e diligência.”²⁷ Também menciona que o dever de diligência tem relação com a qualidade da tomada de decisão pelos administradores, conselheiros, diretores e acionistas controladores.

O autor consigna que as regras de definição destes deveres derivam da Seção 8.30(a) do *Revised Model Business Corporation Act* (RMBCA) de 1984, que estabelecia o dever de o administrador agir com boa-fé e com “a cautela que um homem médio prudente em igual posição teria em circunstâncias semelhantes”²⁸, acreditando que atua de acordo com os melhores interesses para a companhia. Ainda, analisa a alteração promovida no RMBCA:

A seção 8.30 (b) foi alterada para que os administradores quando estiverem informados sobre o processo de tomada de decisão ou quando devotarem atenção suficiente para o processo de supervisão – fossem dispensados de seus

§ 4º O conselho de administração ou a diretoria podem autorizar a prática de atos gratuitos razoáveis em benefício dos empregados ou da comunidade de que participe a empresa, tendo em vista suas responsabilidades

sociais. Apesar de citado por Larissa Quattrini, o referido artigo não foi objeto da presente pesquisa. QUATTRINI, Larissa Teixeira. **Os deveres dos administradores de sociedades anônimas abertas: estudo de casos**. São Paulo: Saraiva, 2014. local 331. *E-book*.

²⁴ QUATTRINI, *loc. cit.*

²⁵ *Ibidem*, local 312.

²⁶ EIZIRIK, Nelson. *A lei das S/A comentada*. 2. ed. rev. e ampl. Artigos 138-205. São Paulo: Quartir Latin, 2015. V. II, p. 115-124.

²⁷ SILVA, Alexandre Couto. **Responsabilidade dos administradores de S/A: business judgment rule**. Rio de Janeiro: Elsevier, 2017. p. 7.

²⁸ *Ibidem*, p. 14.

deveres, quando agirem com o cuidado que uma pessoa em igual posição deveria razoavelmente entender como adequado em circunstâncias similares. O processo de supervisão refere-se ao monitoramento das decisões da administração, enquanto o processo de tomada de decisão refere-se à delegação de determinados assuntos ao Conselho de Administração.

Assim, tem-se imposto elevado padrão de boa-fé no processo de supervisão, fazendo com que os administradores atuem de boa-fé e se esforcem para estabelecer mecanismos que visem a proteger o interesse da companhia de malfeitos de empregados. A tolerância ou a sistemática falha em assegurar informações confiáveis e relatórios adequados constitui uma violação ao dever de diligência.²⁹

Silva descreve as diferentes conceituações e aplicações dadas pelos estados norte-americanos ao dever de diligência, destacando que, no direito brasileiro, o referido dever foi prescrito no art. 153 da Lei das SA. Neste sentido, refere que as dificuldades relativas a este dever surgem quando se observa a série de interpretações que lhe são dadas, no intuito de se tentar obter objetivamente quais comportamentos necessários para que ele seja considerado como cumprido.

Carlos Konrad e Ricardo Lupion examinam o dever de diligência dos administradores e avaliam diversos aspectos deste dever, ponderando que administrar uma sociedade demanda riscos que variam de acordo com a atividade. Nesse sentido, afirmam que o dever pode inibir uma administração mais agressiva e incentivar uma conduta mais conservadora, tendo em vista que o administrador, muitas vezes, precisará tomar decisões rápidas e não poderá depender dos sócios para agir, os quais devem confiar no perfil que escolheram.³⁰

Os autores atentam que o esforço do administrador deve ser maior em casos de temas sensíveis ou relevantes e que:

Os administradores devem explicitar que refletiram, buscaram mais informações e monitoraram o andamento do interesse e avaliação para decidir sobre

²⁹ SILVA, Alexandre Couto. **Responsabilidade dos administradores de S/A: business judgment rule**. Rio de Janeiro: Elsevier, 2017. p. 17.

³⁰ KONRAD, Carlos Bender; LUPION, Ricardo. Dever de diligência dos administradores em tempos de pandemia e impactos da crise nos contratos empresariais. **RJLA, [S.l.]**, ano 7, n. 3, p. 427-456, 2021.

determinado contrato empresarial. O administrador deve inquirir, debater, aprofundar e questionar as informações recebidas sobre a melhor forma de gerir o contrato sensível para a companhia. [...] A decisão e a formalização do seu procedimento se justificam porque o atendimento e prova das obrigações de meio – como o é o dever de agir com diligência – exige formalização para que prejudicados e órgãos reguladores possam averiguar o iter procedimental da Administração, especialmente em contratos sensíveis e relevantes para a companhia.³¹

Ainda, explicam que o ônus de manter registros da gestão e organização é “um meio de conter decisões irresponsáveis, comportamentos oportunistas [...]”³², entre outros prejuízos.

Laís Machado Lucas esclarece que, no direito societário, a análise do cumprimento do dever de diligência só é concretizada a partir das circunstâncias de cada caso e do contexto em que a decisão objeto de avaliação foi tomada, considerando-se o tempo que o administrador tinha para à disposição, as consequenciais possíveis caso tivesse tomado outro caminho ou então demorado para agir. Argumenta que o dever só é visível após a situação, sendo difícil materializá-lo no campo teórico.³³ Reconhece, no entanto, a existência de novos deveres jurídicos que descendem do dever de diligência, que agrupa em: deveres de (i) informar-se e qualificar-se, (ii) de fiscalizar, de investigar e de intervir, (iii) de participar ou de assiduidade, (iv) de bem administrar e de não praticar erros graves. A autora ainda diagnostica um novo dever, que é o de implementar programas de integridade, o que o faz a partir de uma análise sistemática do dever de diligência em conjunto com o art. 7, VIII, da Lei 12.846/2013.³⁴

³¹ *Ibidem*, p. 443.

³² KONRAD, Carlos Bender; LUPION, Ricardo. Dever de diligência dos administradores em tempos de pandemia e impactos da crise nos contratos empresariais. *RJLA*, [S.l.], ano 7, n. 3, p. 427-456, 2021. p. 443.

³³ LUCAS, Laís Machado. **Programas de integridade nas sociedades anônimas**: implementação como conteúdo do dever de diligência dos administradores. Porto Alegre: Livraria do Advogados, 2021. p. 130-131.

³⁴ *Ibidem*, p. 143 e seguintes.

Em livro dedicado a comentar a Lei das Sociedades Anônimas, organizado por Fabio Ulhôa Coelho *et al.*³⁵, Rodrigo Castro e Ana Frazão comentam sobre o art. 153 da Lei das SA.

Na visão de Castro, um administrador que desenvolve e observa os padrões internos de governança da companhia, respeita as esferas de competência de cada um e cumpre as leis, analisa as informações necessárias para as tomadas de decisão e busca atingir os melhores interesses e resultados para a companhia é cumpridor do dever de diligência.

Frazão elencou condutas objetivas que garantem o cumprimento do referido dever em dez itens: (i) dever de informação para a tomada de decisões, (ii) competência profissional para o cargo, (iii) perícia compatível com as funções desempenhadas, (iv) conhecimento sobre os negócios da companhia, (v) estruturação de uma organização empresarial compatível com o porte e risco assumido pela cia, (vi) estruturação de um sistema de vigilância, supervisão e investigação, (vii) dever de intervir diante de problemas ou ameaças, (viii) respeito às obrigações legais, sem descumprimentos "calculados", (ix) presença nas reuniões dos colegiados que é parte e (x) utilização de processos decisórios justos, adequados e razoáveis.

Frazão descreve o dever de diligência como uma metodologia decisória, que serve para avaliar o processo de tomada de decisão e não seus resultados práticos, propriamente ditos. Assim, o dever de diligência pode ter existido mesmo que os resultados práticos de determinada decisão tenham sido negativos (e o inverso também), pois o foco na atribuição da responsabilidade não deve ser a dimensão do dano causado, mas avaliar se todas as cautelas possíveis foram tomadas para evitá-lo ou mitigá-lo. Ou seja, "o foco do dever de diligência não é o que foi decidido, mas sim como foi decidido."³⁶

³⁵ COELHO, Fábio Ulhoa (Coord.) *et al.* **Lei das sociedades anônimas comentada**. Rio de Janeiro: Forense, 2021. p. 873 e segs.

³⁶ FRAZÃO, Ana. Responsabilidade civil dos administradores de sociedades empresárias por decisões tomadas com base em sistemas de inteligência artificial. *In*: FRAZÃO, Ana; MULHOLLAND, Caitlin (Coord.). **Inteligência artificial e Direito: ética, regulação e responsabilidade**. 2. ed. rev. atual. e ampl. São Paulo: Thomson Reuters Brasil, 2020. p. 81-94. *E-book*.

Conforme a autora, agir bem-informado é o núcleo essencial deste dever. Contudo, em razão da fluidez de seu conceito, reporta que discussões recentes propõem a ampliação do núcleo do dever para integrar a adoção de programas de *compliance*³⁷ como conduta elementar.

Frazão complementa que o *compliance* reforça a dimensão organizacional do dever de diligência, já que cria “adequados sistemas de vigilância, supervisão e investigação sobre as atividades da sociedade, de modo a assegurar o respeito às obrigações legais e possibilitar a intervenção adequada diante da identificação de problemas e ameaças.”³⁸ Isso porque estes sistemas proporcionam mecanismos que antecipam e até evitam riscos, danos e/ou ilícitos, cujo porte e estrutura devem ser analisados sob o contexto em que a companhia está inserida. Reconhece ser importante que o dever de diligência seja assegurado por estratégias de governança que estipulem procedimentos para garantir cuidado e zelo nas tomadas de decisão.

Além das condutas práticas acima relacionadas, Lupion atenta para a postura esperada de um administrador diligente:

O dever de diligência envolve tempo dedicado à função. Participação ativa - agir, inquirir e avaliar alternativas e consequências. Compartilhar conhecimento e discutir com os demais administradores. Assessorar-se com especialistas.

[...]

Daí, então, surgem os chamados da regra da *standards* decisão empresarial, que representam um mecanismo formal e seguro para impor aos administradores um agir consciente e direcionado à obtenção dos melhores resultados para a empresa. Essa regra, como já referido, preocupa-se apenas com o processo que levou à decisão e não com o seu mérito.

[...]

Além dos já assinalados da regra da *standards* decisão empresarial, comuns ao dever de diligência, também é exigível do administrador que investigue se as

³⁷ Descrito pela autora como o conjunto de ações a serem adotadas no ambiente corporativo para que se reforce a anuência da empresa à legislação vigente, de modo a prevenir a ocorrência de infrações ou, já tendo ocorrido o ilícito, propiciar o imediato retorno ao contexto de normalidade e legalidade.

³⁸ FRAZÃO, Ana. Dever de diligência: Novas perspectivas em face de programas de *compliance* e de atingimento de metas. **JOTA**. 15 fev. 2017. Disponível em: <http://jota.info/colunas/constituicao-empresa-e-mercado/dever-de-diligencia-15022017>. Acesso em: 17 out. 2022.

informações que lhe são passadas são confiáveis, suficientes e corretas. Isso significa que, não basta que o administrador da companhia informe-se sobre as situações que envolvem a decisão a ser tomada, é necessário que exista uma análise crítica a fim de detectar eventuais e potenciais problemas que possam ocorrer na empresa. O dever de investigar somente se impõe diante das circunstâncias específicas que levem os administradores a acreditarem que existe algum tipo de risco, mesmo que potencial, para a companhia, ou seja, quando os administradores se encontrarem diante de situações que os coloquem em “estado de alerta”.³⁹

Todo o exame feito em relação à conceituação dada pelas autoras e pelos autores acima referidos revela que a doutrina é plural na conceituação do dever de diligência. Alguns autores entendem que o padrão de conduta adotado é o do bom pai de família, enquanto outros entendem que é o homem de negócios com capacidades técnicas para tanto.

Apesar disso, há um relativo consenso sobre (i) ser um dever de meio (e não de resultado), (ii) ser flexível para que se ajuste conforme o contexto em que a decisão foi tomada, a partir do caso concreto e (iii) ter evoluído com a passagem do tempo, adequando-se as exigências devidas às práticas de mercado e porte da empresa.

Por outro lado, esta pluralidade enseja diversas interpretações sobre quais são as posturas que um(a) administrador(a) deve adotar a fim de que seja futuramente reconhecido que todas as diligências necessárias foram seguidas para, ao menos, evitar os danos ocasionados a partir de tomadas de decisão arriscadas ou prejudiciais à empresa.

Dentre os autores estudados, foi possível mapear diversas ações (que poderiam ser classificadas como subdeveres) imprescindíveis de serem seguidas antes de uma tomada de decisão empresarial. Os subdeveres mais citados são: dever de se informar, dever de se qualificar, dever de bem administrar, dever de investigar, dever de vigiar/vigilância, dever de supervisão/monitoramento, dever de

³⁹ LUPION, Ricardo. Contratos empresariais e o mundo BANI: Brittle, Anxious, Nonlinear and Incomprehensible. *Revista Jurídica Luso-Brasileira - RJLB*, [S.l.], ano 8, n. 6, p. 1942-1944, 2022.

intervir, dever de não praticar erros graves, dever de participação/assiduidade e outros deveres envolvidos, os quais serão examinados, um a um, a seguir.

3.2 Dever de se informar

Elencado por Frazão como o núcleo elementar do dever de diligência, o dever de se informar é visto como condição mínima para uma tomada de decisão diligente e é a partir deste pressuposto que as diligências se desdobram. Tal dever impõe ao administrador a necessidade de obter o máximo de informações possíveis “a respeito da matéria relacionada com as decisões que pretendam adotar no sentido de cumprirem o seu dever de bem administrar.”⁴⁰

Parente aduz que:

Os administradores devem, por fim, se manter permanentemente informados sobre os procedimentos, programas e outras técnicas que os auxiliem na função de supervisão das atividades desenvolvidas no âmbito da companhia.

[...]

Verifica-se, portanto, que os administradores têm o dever de obter todas as informações razoavelmente disponíveis antes de tomarem quaisquer decisões relativas aos negócios da companhia.

Com relação a este dever, o desafio é entender qual o grau de informação exigível dos administradores e como saber se cumpriram de forma adequada o seu dever de se informar. A avaliação dependerá das circunstâncias de cada caso, devendo ser considerado (i) o conhecimento prévio sobre o assunto a ser definido, (ii) o tempo disponível para a tomada de decisão, (iii) os custos de um estudo pormenorizado sobre o tema, (iv) e a importância da decisão para a companhia.⁴¹

Já Brigagão destaca que o seu cumprimento consiste na “obrigação dos dirigentes em buscar informações quando da tomada de decisão negocial”, bem como de “estudar as operações da empresa de um modo geral, colhendo as devidas

⁴⁰ PARENTE, Flávia. **O dever de diligência dos administradores de sociedades anônimas**. Rio de Janeiro: Renovar, 2005. p. 111.

⁴¹ *Ibidem*.

informações para que seja capaz de tomar uma decisão diligente."⁴² Brigagão cita que a importância do dever de se informar foi maior evidenciada a partir do julgamento da venda da empresa Trans Union, no caso conhecido como Smith vs. Van Gorkom, e destaca que a decisão condenatória entendeu que o conselho não se informou substancialmente acerca do valor efetivo das ações da Trans Union, não se informando acerca de todos os dados razoavelmente disponíveis e relevantes para a decisão que foi tomada.⁴³

Para Campos, o dever de se informar: "significa [...] que o administrador deve procurar se informar sobre os negócios da companhia de uma maneira geral e obter os dados necessários para uma diligente tomada de decisão." Ainda, "[...] pressupõe que se obtenham as informações adequadas para se avaliar os riscos do negócio, as vantagens e desvantagens que a decisão a ser tomada envolve."⁴⁴

Ribeiro também destaca que, no conteúdo do dever de diligência, encontra-se o dever de informar-se, que se configura em "adquirir conhecimento sobre a sociedade e o ramo de negócio."⁴⁵ Entende que a diligência exige o dever de buscar informações, tendo em vista que as decisões dos gestores devem estar de acordo com os deveres legais, exigindo que o administrador busque:

as informações necessárias para a tomada de decisões [...], tais como [...] conhecimento da situação de mercado, política e produtos da empresa, dificuldades e propostas de soluções, [...] considerando-se o conhecimento anterior da matéria, tempo para a tomada de decisão, sua importância para a empresa, possibilidade ou não de perda de oportunidade, de negócio, custos para obtenção de informações, vantagens e desvantagens da medida proposta entre outras". [...]

⁴² BRIGAGÃO, Pedro Henrique Castello. **Dever de diligência dos administradores de sociedades anônimas**. Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Direito. Rio de Janeiro, 2012. Disponível em: https://www.puc-rio.br/ensinopesq/ccpg/pibic/2012/resumos_pdf.pdf. Acesso em: 17 out. 2022. p. 3.

⁴³ BRIGAGÃO, *loc. cit.*

⁴⁴ CAMPOS, Luiz Antonio de Sampaio. Deveres e responsabilidades. In: LAMY FILHO, Alfredo; PEDREIRA, José Luiz Bulhões (Coord.). **Direito das companhias**. São Paulo: Forense, 2009. p. 1101-1105.

⁴⁵ RIBEIRO, Renato Ventura. **Dever de diligência dos administradores de sociedades**. São Paulo: Quartier Latin, 2006. p. 224.

Se a operação gera prejuízos, por ter sido tomada com falta de informações ou informações inadequadas, tem-se caracterizada a culpa dos administradores, por falta do dever de diligência.⁴⁶

O autor também refere sobre como um procedimento que conduza fluxos amparados por boa informação podem ser causas de isenção de responsabilidade, tendo em vista que "decisões tomadas de boa-fé, com base em adequada formação de juízo, através de informações fundamentadas e confiáveis, e cuja racionalidade indica ser o melhor para a companhia afastam a presunção de culpa."⁴⁷

O artigo de Konrad e Lupion também aborda o quanto a administração de uma sociedade deve obter todas as informações necessárias antes de tomadas de decisão estratégicas para empresa, mesmo em momentos de crise, devendo-se contar, inclusive, com especialistas no tema; a obra também reforça como a formalização do procedimento que levou a uma tomada de decisão cautelosa é relevante para o cumprimento do dever de diligência:

Em relação ao dever de se informar, extrai-se que a Administração deve obter todas as informações possíveis e razoáveis a respeito da matéria relacionada com as decisões que pretendem adotar.⁴⁸ [...] os administradores devem buscar mais dados para uma decisão informada e refletida, podendo socorrer-se de profissionais da área jurídica, corpo gerencial, consultores externos, profissionais da área econômica e contábil para definir como agir em relação a um instrumento obrigacional relevante da empresa. Em qualquer caso, a necessidade de adequada informação como elemento essencial na regra da decisão empresarial e como concretização do dever de diligência não pode ser afastada pela alegação da urgência da decisão empresarial. Isto é, decisões urgentes demandam, talvez com maior rigor, o atendimento dos pressupostos dos deveres dos administradores da companhia aberta. [...].⁴⁹

⁴⁶ *Ibidem*, p. 227-228.

⁴⁷ *Ibidem*, p. 232.

⁴⁸ KONRAD, Carlos Bender; LUPION, Ricardo. Dever de diligência dos administradores em tempos de pandemia e impactos da crise nos contratos empresariais. **RJLA**, [S.l.], ano 7, n. 3, p. 427-456, 2021. p. 446.

⁴⁹ KONRAD, Carlos Bender; LUPION, Ricardo. Dever de diligência dos administradores em tempos de pandemia e impactos da crise nos contratos empresariais. **RJLA**, [S.l.], ano 7, n. 3, p. 427-456, 2021. p. 447.

[...]

a formalização do iter, isto é, as informações, documentos, cenários, opiniões de experts, projeções etc. que levaram à essa conclusão devem ser suficiente e adequadamente registradas no ato decisório (ata da reunião) e ficarem disponíveis para avaliação posterior pela companhia, por acionistas, por órgãos reguladores ou por julgadores e árbitros. Neste caso, ainda assim, permanecem íntegros e acentuados os deveres fiduciários.⁵⁰

Inclusive, do dever de se informar, há um direito implícito de se fiar no assessoramento prestado por terceiros (*right do rely*), eis que não é possível exigir dos administradores que tenham conhecimentos técnicos suficientes para lidar com todos os tipos de assuntos e especialidades. Sendo assim, ao administrador é reconhecida a faculdade de consultar um especialista na área que precisa ser analisada para a tomada de decisão. Deste modo, ainda que a opinião do *expert* consultado gere um grande prejuízo para a empresa, o cumprimento do dever de diligência, especificamente do dever de se informar, poderá ser reconhecido pela Comissão de Valores Mobiliários⁵¹, quando cumpridos alguns requisitos:

a) **Qualificação do assessor consultado:** requer que o administrador certifique a experiência do assessor no tema a ser consultado, bem como a ausência de conflito de interesse para emitir um parecer sobre o assunto;

b) **Divulgação de informações:** exige que os administradores forneçam todas as informações e subsídios necessários para que o assessor possa analisar a questão objeto de consulta, considerando o contexto e as variáveis existentes;

c) **Efetiva confiança:** o fato de consultar um assessor não exime o administrador de fazer uma análise crítica sobre as orientações emitidas. Se entender necessário, inclusive é indicado solicitar mais esclarecimentos, detalhamentos, ou até mesmo obter mais de uma opinião;

⁵⁰ *Ibidem*, p. 450.

⁵¹ Conforme emblemático julgamento sobre o tema: BRASIL. Comissão de Valores Mobiliários. **Processo Administrativo Sancionador CVM nº RJ2014/8013**. Voto Diretor Relator Gustavo Machado Gonzales. Fls. 45 e 46. 2018. Disponível em: https://conteudo.cvm.gov.br/sancionadores/sancionador/2018/20180731_PAS_RJ20148013.html. Acesso em: 09 jan. 2023.

d) **Nexo de causalidade:** o direito de se fiar no assessoramento prestado por terceiros só pode ser invocado para o caso de a opinião técnica do assessor ter sido adotada pelos administradores. Ou seja, caso os administradores optem por desconsiderar a opinião técnica do assessor, ainda que embasados por justificativas, não poderão invocar o *right to rely* – o que não impede que demonstrem eventual diligência por outras cautelas adotadas;

e) **Sinais de alerta:** como já mencionado anteriormente, o administrador não esgota o cumprimento do dever de se informar por simplesmente obter a orientação de um especialista. Ele precisa avaliá-la com cautela e considerar todos os riscos. Isto exige que esteja atento a eventuais sinais de alerta que podem/devem ser considerados em sua tomada de decisão;

f) **Boa-fé:** a existência de boa-fé é premissa básica para o reconhecimento do referido direito, sob pena de a contratação de assessores mascarar condutas desonestas. Caso se comprove que o administrador agiu de má-fé, não há como defender sua diligência com base em assessoramento externo.

A conceituação do dever de se informar revela que ele é o centro do dever de diligência. Isso porque é a partir das informações obtidas que será possível avaliar quais as cautelas a serem adotadas, tais como a necessidade de investigar, de se qualificar sobre o tema, de supervisionar e monitorar. A informação é a base da avaliação de risco e somente de posse dela é que se pode cumprir os demais subdeveres. Sem cumprir com o dever de se informar, não há como ser diligente, o que evidencia que ele efetivamente é um pressuposto para o cumprimento do dever de diligência. Registra-se que Quattrini, Silva, Yazbek e Lucas, também dão o devido destaque a este subdever em suas obras.

3.3 Dever de se qualificar

Sobre o dever de se qualificar, Parente esclarece que se trata do nível de profissionalismo/profissionalização do administrador, suas capacidades, competências e qualificações técnicas especializadas para o exercício da função. Ela ainda atenta que

não podem ser exigidos do administrador a perícia e os conhecimentos técnicos em todas as áreas relativas às atividades desenvolvidas pela sociedade, contudo, se ele for diligente, poderá obter junto a colaboradores e até consultores externos a condução do negócio de maneira satisfatória.⁵²

Ribeiro também destaca que, no conteúdo do dever de diligência, encontra-se o dever de qualificar-se, para o qual atribui a mesma definição do dever de informar-se, qual seja: “adquirir conhecimento sobre a sociedade e o ramo de negócio.”⁵³

Sobre o dever de se qualificar, Brigagão confere a ele duas necessidades: a primeira deve ser analisada *a priori*, ou seja, uma qualificação profissional mínima deve ser necessária para que o administrador exerça a incumbência, devendo recusá-la caso não se sinta minimamente qualificado. A segunda necessidade é a de se manter qualificado durante todo período em que exercer o cargo, aprimorando seus conhecimentos técnicos e se atualizando.

Campos⁵⁴ avalia a exigência de possuir profissionalidade e perícia para que um administrador seja considerado diligente – exigência que, conforme sua interpretação, não encontra guarida na legislação de regência, eis que a lei das sociedades anônimas não exige habilitação técnica e profissional, bastando que o sujeito possua meios e energias para exercer a função de administrador. Em seu entendimento, compreende que a avaliação sobre a aptidão de determinada pessoa ao cargo é do acionista controlador que o elege e dele próprio, que deve negar caso não se sinta competente para tanto, não sendo indicado que se aventure em situações que não conheça, devendo se socorrer do auxílio de *experts*, nos termos já relatados anteriormente.

Este subdever realmente merece ser avaliado com cautela. Isso porque, além de a legislação de regência não impor nenhuma competência técnica ou especializada, como bem afirmou Campos, há inúmeros casos de administradores de grandes empresas que as conduziram com diligência e sucesso em seus

⁵² PARENTE, Flávia. **O dever de diligência dos administradores de sociedades anônimas**. Rio de Janeiro: Renovar, 2005. p. 105.

⁵³ RIBEIRO, Renato Ventura. **Dever de diligência dos administradores de sociedades**. São Paulo: Quartier Latin, 2006. p. 224.

⁵⁴ CAMPOS, Luiz Antonio de Sampaio. Deveres e responsabilidades. *In*: LAMY FILHO, Alfredo; PEDREIRA, José Luiz Bulhões (Coord.). **Direito das companhias**. São Paulo: Forense, 2009.

negócios, mesmo sem formação específica ou experiência prévia.⁵⁵ Entretanto, como advertiu Brigagão, durante o exercício da atividade empresarial, é importante que o administrador se qualifique profissionalmente, seja se envolvendo nas rotinas da empresa e observando o mercado em que está inserido, seja buscando cursos e formações em sua área de atuação ou cercado-se de equipe qualificada que lhe assessorar.

3.4 Dever de bem administrar

Com relação ao dever de bem administrar, Parente o designa como um dever-poder, que comporta a gestão e a representação da companhia, ou seja, “praticar todos os atos necessários à plena execução da vontade social, visando realizar o objeto social da companhia.”⁵⁶ A autora ainda destaca que os administradores têm a discricionariedade de eleger os meios mais adequados para consecução do objeto social, o que devem fazer de maneira diligente.⁵⁷

Brigagão define o dever de bem administrar como aquele que prescreve que o administrador realize todos os atos necessários na persecução do objeto social, dentro de sua discricionariedade, mas sempre nos limites impostos pela lei e pelo estatuto.⁵⁸ Lucas comenta acerca da subjetividade do dever de bem administrar, sendo impossível a sua exata definição, certo apenas que este não depende do sucesso ou insucesso de uma decisão negocial.⁵⁹

⁵⁵ Como exemplos que podem ser citados, temos Nestor de Paula e Geraldo Rufino, que fundaram e administraram grandes companhias mesmo sem nenhuma formação técnica. Veja-se, nesse particular: MORRE fundador da marca Azaleia Calçados. In: AGÊNCIA ESTADO. Folha de Londrina. 24 jan. 2004. Disponível em: <https://www.folhadelondrina.com.br/geral/morre-fundador-da-marca-azaleia-de-calcados-477027.html> Acesso em: 17 out. 2022; GERALDO RUFINO: o ex-catador de latinhas que criou uma das maiores empresas de reciclagem de caminhões. In: INFOMONEY, 2023. Disponível em: <https://www.infomoney.com.br/perfil/geraldo-rufino/>. Acesso em: 16 out. 2022.

⁵⁶ PARENTE, Flávia. **O dever de diligência dos administradores de sociedades anônimas**. Rio de Janeiro: Renovar, 2005. p. 109.

⁵⁷ *Ibidem*, p. 110.

⁵⁸ BRIGAGÃO, Pedro Henrique Castello. **Dever de diligência dos administradores de sociedades anônimas**. Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Direito. Rio de Janeiro, 2012. Disponível em: https://www.puc-rio.br/ensinopesq/ccpg/pibic/relatorio_resumo2012/.pdf. Acesso em: 17 out. 2022. p. 6.

⁵⁹ LUCAS, Laís Machado. **Programas de integridade nas sociedades anônimas: implementação como conteúdo do dever de diligência dos administradores**. Porto Alegre: Livraria do Advogados, 2021. p. 150.

O que se denota é que o reconhecimento de uma boa administração requer uma análise de um conjunto de atitudes tomadas pelo administrador, sem, contudo, obter consenso sobre quais são estas atitudes. Apesar de os autores citados defenderem que a boa administração é intrínseca ao dever de diligência, esse subdever parece mais pertencer a uma etapa posterior: somente quando o administrador atuar com cautelas e diligência é que será possível definir, *a posteriori*, se a administração foi, ou não, bem-feita.

3.5 Dever de investigar

Parente destaca que o dever de investigar possui dois aspectos: examinar de forma crítica a informação e vigiar a detecção de incongruências ou situações “estranhas” que possam gerar danos ou prejuízos à sociedade. A investigação é exigida do administrador especialmente quando houver suspeitas de riscos, mesmo que potenciais – o que importa dizer que, caso não haja motivos para desconfiar das informações analisadas, não se exige do administrador uma investigação. Assim como o dever de se informar, o dever de investigar também considera a oportunidade e conveniência de o administrador promover uma investigação, considerando-se tempo e custos despendidos para tanto.⁶⁰

Lupion e Konrad também examinam este subdever:

O dever de vigiar ou investigar exige que os administradores analisem criticamente as informações que lhes foram fornecidas para detectar possíveis problemas futuros. Assim, exige-se que a administração discuta criticamente a informação e que avalie se as disponíveis são suficientes para uma decisão ou se devem ser complementadas.

[...] o dever de agir com diligência, é adimplido por um fazer, uma conduta não especificamente voltada para um resultado, de modo que a Comissão de Valores

⁶⁰ CAMPOS, Luiz Antonio de Sampaio. Deveres e responsabilidades. *In*: LAMY FILHO, Alfredo; PEDREIRA, José Luiz Bulhões (Coord.). **Direito das companhias**. São Paulo: Forense, 2009. p. 1108.

Mobiliários destaca que o foco para verificar o seu adimplemento é no aspecto procedimental.⁶¹

Campos compreende este subdever como uma decorrência do dever de se informar, que impõe a obrigação de o administrador "investigar fatos de que tomem conhecimento ou que despertem algum interesse em particular".⁶² Considerando que a viabilidade de instaurar uma investigação pressupõe diversos fatores, o autor se apoia nos *Principles of the American Law Institute* e recomenda que o dever de investigar surge quando houver indícios razoáveis de riscos, ou seja, "circunstâncias (que) indiquem a existência de irregularidade."⁶³

Brigagão também cita o dever de investigar (decorrente do dever de vigilância, mencionado a seguir) como aquilo que deve ser feito sempre que o administrador supor que as informações obtidas estão equivocadas, incompletas, superficiais, imprecisas etc. Destaca que, para promover uma investigação, o administrador deve possuir indícios de que há possibilidade de algum problema, analisando se é razoável despender tempo e valores demandados para uma investigação, o que precisa ser avaliado sempre que surgirem sinais de alerta ou *red flags*.⁶⁴

⁶¹ KONRAD, Carlos Bender; LUPION, Ricardo. Dever de diligência dos administradores em tempos de pandemia e impactos da crise nos contratos empresariais. *RJLA*, [S.l.], ano 7, n. 3, p. 427-456, 2021. p. 448.

⁶² CAMPOS, *loc. cit.*

⁶³ *Ibidem*, p. 1110.

⁶⁴ Segundo Pedro Henrique Castello Brigagão: As *red flags* caracterizam-se por circunstâncias nas quais o administrador entende estar diante de eventos que coloquem os negócios da companhia em situação de grave e manifesto risco, "situações que o deixem em estado de alerta" Assim, ao tempo em que o administrador suspeitar ou identificar que a companhia está diante de situações que possam causar problemas relevantes no desenvolvimento dos negócios da companhia, deverá o mesmo agir e promover investigação minuciosa em relação a essas eventuais e possíveis dificuldades. Então, as *red flags* funcionam como verdadeiros sinais de alerta para o administrador. Ele pode, com isso, confiar nas informações que são disponibilizadas por auditores e profissionais competentes referentes à situação dos negócios da companhia, mas essa confiança não é inquebrável. Isto é, se o administrador suspeitar que os dados que lhes são fornecidos aparentem ser incompatíveis com o real estado da companhia, caracterizando a ocorrência dos sinais de alerta, deverá empenhar diligente investigação a fim de resolver essa incompatibilidade. BRIGAGÃO, Pedro Henrique Castello. **Dever de diligência dos administradores de sociedades anônimas**. Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Direito. Rio de Janeiro, 2012. Disponível em: https://www.pucrio.br/ensinopesq/ccpg/pibic/relatorio_resumo2012/resumos_pdf/ccs/DIR/Pedro%20Henrique%20Castello%20Brigag%20E3o.pdf. Acesso em: 17 out. 2022. p. 5.

Sobre este dever, Quattrini analisou o caso envolvendo a condenação da empresa de resseguro Pritchard & Baird Intermediaries Corporation ("P&B"), fundada por Charles Pritchard. Após o falecimento de Charles, as ações da empresa foram repassadas para sua viúva Lilian (48%) e o restante entre seus dois filhos, Charles Jr. e William, os quais foram responsabilizados pessoalmente por negligenciarem o dever de diligência na condução dos negócios empresariais – atitude que levou à falência da empresa.⁶⁵

O motivo da condenação foi o fato de que Charles Jr. e William retiravam empréstimos de fundos de custódia que deveriam servir aos clientes da empresa de resseguros, o que configurou apropriação indébita por parte dos acionistas. Sua mãe, Lilian, não tomava conhecimento das demonstrações financeiras e das atividades empresariais e, conforme decisão, não fez esforços para assegurar que a política e prática da sociedade com relação aos saques de fundos respeitasse a legislação vigente e os costumes do setor, sem atentar para seus deveres como membro do Conselho de Administração. Mesmo com debilitado estado de saúde e mental – o que se agravou após o falecimento de seu marido –, Lilian foi considerada negligente na assunção de suas obrigações. Segundo Quattrini, o julgamento considerou que "o dever de diligência é inexorável ao cumprimento de seus deveres, de sorte a não se admitir defesa respaldada em pretensa falta de conhecimento para seu cumprimento."⁶⁶

Com base neste entendimento, a Corte de New Jersey estabeleceu um padrão de comportamento esperado de um membro de Conselho de Administração, qual seja:

- (i) Busca pelo conhecimento mínimo necessário para exercer deveres para ele exigidos quando entender pela insuficiência de sua experiência;
- (ii) Obrigação contínua e ininterrupta de manter-se informado acerca das atividades da Companhia;
- (iii) Acompanhamento das condutas da Companhia, não sendo admitida alegação de ignorância de má conduta desta, principalmente para se esquivar da

⁶⁵ QUATTRINI, Larissa Teixeira. **Os deveres dos administradores de sociedades anônimas abertas: estudo de casos**. São Paulo: Saraiva, 2014. *E-book*.

⁶⁶ *Ibidem*, local 800 de 5659.

verificação de determinado fato, que seria evidente por meio de simples verificação;

(iv) Ainda que aos órgãos superiores da Administração não caiba uma inspeção detalhada das atividades diárias da Companhia, cabe-lhes uma fiscalização geral dos negócios e das políticas empresariais; e

(v) Não se espera que os conselheiros façam completa auditoria dos livros da Companhia, porém devem manter familiaridade com sua posição financeira, mediante revisão das demonstrações financeiras.⁶⁷

A Corte ressaltou que o Conselho de Administração pode ser isento de responsabilidade quando, de boa-fé, toma decisões baseadas na opinião de consultores da Companhia, ou em relatórios escritos, preparados por contadores independentes ou em demonstrações financeiras, livros contábeis etc. No caso concreto, a Corte entendeu que Lilian, como membro do Conselho de Administração, tinha o dever de averiguar e promover uma investigação mais apurada sobre os “empréstimos para os acionistas” retirados dos fundos de custódia. Isso porque, tratando de uma empresa de resseguros, esta tem uma responsabilidade adicional com relação às demais empresas, pois seus deveres não são somente com relação aos acionistas, mas também relacionados a sua função de depositária de empresas cedentes e resseguradoras.

Em que pese seja criticável o resultado do julgamento em razão do estado de saúde específico de Lilian, que já estava acometida por doença que lhe incapacitava de exercer plenamente suas atividades, a conclusão do julgado ratifica a necessidade de maior avaliação das informações submetidas ao conhecimento dos administradores. De fato, os administradores, para serem considerados diligentes, devem investigar sempre que houver sinais de alerta de potenciais danos à empresa, o que não significa que, uma vez devidamente investigado, eles não possuam discricionariedade para seguir na tomada de decisão independentemente das *red flags* que foram investigadas e posteriormente superadas.

Ademais, é importante ressaltar que o formato, o tempo e o custo despendido para esta investigação são também definidos pela autonomia do administrador

⁶⁷ *Ibidem*, local 813 de 5659.

conforme o porte e a disponibilidade da companhia. O que é relevante é que o administrador dedique espaço para que os alertas sejam devidamente apurados e compreendidos na tomada de decisão.

3.6 Dever de vigiar/vigilância

No que toca ao dever de vigiar, Parente ressalta a obrigação de fiscalização constante, supervisionando e monitorando os atos da companhia praticados por suas equipes, mantendo o controle do que ocorre ou do que deixa de ocorrer. Trata-se de um acompanhamento geral, que precisará ser mais acurado caso se trate um assunto estratégico para a sociedade.

Destaca a autora que está inserida, no dever de vigiar ou monitorar, a obrigação de:

Os administradores verificarem se a companhia possui programas impondo a seus funcionários e subordinados o exato cumprimento de normas estatutárias, legais e regulamentares, que lhes forem aplicáveis ("*compliance programs*"), cabendo-lhes, ainda, assegurarem-se de que tais programas estão sendo seguidos.⁶⁸

Sobre este subdever, Quattrini também constata que uma das razões para a condenação mencionada no tópico anterior foi porque a viúva Lilian também não supervisionou e nem monitorou a empresa administrada por seus filhos:

[...] entendeu a Corte que, enquanto conselheira de uma substancial companhia corretora de resseguro, Sra. Pritchard teria como obrigação: (a) a ciência, de que a companhia recebia, anualmente, milhões de dólares de fundos de prêmios e perdas (*loss and premium funds*), que eram mantidos em custódia para empresas cedentes e resseguradoras; (b) obtenção e estudo das demonstrações financeiras, sendo certo que, embora tivesse o direito de confiar em tais documentos, elaborados de acordo com a N.J.S.A. 14A:6-14, tal confiança não

⁶⁸ PARENTE, Flávia. **O dever de diligência dos administradores de sociedades anônimas**. Rio de Janeiro: Renovar, 2005. p. 131.

justificaria sua conduta, já que tais documentos divulgavam claramente a apropriação indevida dos fundos de fideicomisso pelos seus filhos.

[...]

Em suma, entendeu a Corte que a Sra. Pritchard teria como obrigação conhecer e supervisionar minimamente os negócios da Pritchard & Baird, mediante a análise das demonstrações financeiras e a elaboração de razoáveis tentativas de detectar e prevenir condutas e/ou atos ilegais dos demais conselheiros e diretores. Por conseguinte, a Sra. Pritchard violou o seu dever de proteger os clientes da Pritchard & Baird contra políticas e práticas que poderiam resultar em apropriações indevidas de dinheiro confiado à Companhia.⁶⁹

Em sua análise crítica sobre o caso, Quattrini destaca que este julgamento foi o primeiro a definir, em linhas gerais, qual o padrão de comportamento esperado de um conselheiro, com destaque para quatro itens: "(i) ter conhecimentos mínimos necessários para o exercício do cargo; (ii) manter-se informado acerca das atividades da empresa, mediante participação em reuniões; (iii) realizar o acompanhamento dos negócios sociais; e (iv) fazer fiscalização geral dos negócios."⁷⁰ A autora também ressalva que a Corte de New Jersey destacou que não se espera dos conselheiros uma inspeção minuciosa das atividades diárias de gestão, tampouco dos livros da companhia, mas que, apesar disso, deixa claro que o "o dever de monitoramento das atividades compõe o dever de diligência."⁷¹ A pesquisadora também destacou que o padrão mínimo de conduta estabelecido pela Corte de News Jersey é plenamente aplicável ao direito brasileiro, conforme os princípios estabelecidos pela Lei 6.404/76.

Silva menciona que se exige do administrador postura proativa na fiscalização da companhia, inclusive com relação à observância dos demais princípios de governança corporativa.⁷² Acresce ainda que, mais do que uma supervisão, o cumprimento do subdever envolve um sistema de monitoramento das atividades e

⁶⁹ QUATTRINI, Larissa Teixeira. **Os deveres dos administradores de sociedades anônimas abertas: estudo de casos**. São Paulo: Saraiva, 2014. local 845 e 860 de 5659. *E-book*.

⁷⁰ *Ibidem*, local 896 de 5659.

⁷¹ *Ibidem*, local 896 de 5659.

⁷² SILVA, Alexandre Couto. **Responsabilidade dos administradores de S/A: business judgment rule**. Rio de Janeiro: Elsevier, 2017. p. 7.

das informações, com o intuito de assegurar que os interesses e direitos dos acionistas não estejam sendo desrespeitados em instâncias inferiores das empresas. O autor destaca também que o monitoramento interno da companhia deve ser permanente, revisando-se inclusive a estrutura da companhia, de modo a assegurar diretrizes claras sobre as funções e atribuições da administração da companhia. Informa que o sistema deve ser divulgado de modo que os acionistas e o mercado possam examinar sua adequação.

Campos, por sua vez, descreve que o dever de vigilância consiste em:

[...] uma vigilância geral, no sentido de monitorar o andamento dos negócios e a execução das deliberações e decisões tomadas. O dever de vigiar – ou monitorar, conforme a doutrina norte-americana – o desenvolvimento das atividades sociais não impõe a obrigação de examinar detalhadamente todas as operações da sociedade.

[...]

Isso, contudo, não exige os administradores de acompanharem mais proximamente operações específicas cuja relevância assim o recomende.⁷³

Ribeiro afirma que o dever de vigiar significa:

supervisão e fiscalização dos assuntos e pessoas no âmbito da competência do administrador. Abrange todos os assuntos relacionados à competência do administrador, com delegação ou não de poderes, incluindo atos de outros administradores, subordinados, terceiros responsáveis pela gestão social, procuradores e das sociedades controladas.

[...]

O tempo, esforço e custos a serem despendidos na vigilância dependem de cada situação. Os assuntos mais importantes ou delicados exigem vigilância mais cuidadosa.⁷⁴

⁷³ CAMPOS, Luiz Antonio de Sampaio. Deveres e responsabilidades. *In*: LAMY FILHO, Alfredo; PEDREIRA, José Luiz Bulhões (Coord.). **Direito das companhias**. São Paulo: Forense, 2009. p. 1107-1108.

⁷⁴ RIBEIRO, Renato Ventura. **Dever de diligência dos administradores de sociedades**. São Paulo: Quartier Latin, 2006. p. 224.

O exame desse subdever demonstra a necessidade de o administrador estar presente nas rotinas da companhia e de suas equipes. A depender do porte da empresa e da formação técnica do administrador, este subdever pode e deve ser delegado a funcionários capacitados para o exercício das atividades, a fim de garantir o regular desenvolvimento da empresa. Aqui, também se denota a necessidade do estabelecimento de fluxos e procedimentos que auxiliem nesta vigília, como por exemplo reuniões de *follow up* e *feedbacks* periódicos.

3.7 Dever de supervisão/monitoramento

Daniele de Lima Oliveira comenta a inspiração do modelo brasileiro no *Revised Model Business Corporation Act* (RMBCA) de 1984 e afirma que a modificação nele feita em 1999 impõe o monitoramento das decisões da administração, com elevado padrão de boa-fé e exige o estabelecimento de mecanismos que busquem proteger os interesses da companhia em relação à atuação dos seus empregados. Ato seguinte, a autora destaca sobre o dever de supervisão dos administradores sobre os empregados e demais subordinados da companhia, exemplificando que as falhas no dever de diligência podem decorrer de atos comissivos ou mesmo omissivos, mais comumente vinculados a atos de negligência na supervisão de seus subordinados.⁷⁵

Yazbek também menciona que, na atualidade, o dever de diligência evoluiu para uma espécie de dever de monitoramento das atividades da administração.⁷⁶

Como se vê, tais autores mencionam o dever de supervisão/monitoramento como sendo uma nova categoria autônoma e independente. Contudo, em sua conceituação, percebe-se que este subdever é bem próximo do subdever de vigilância acima descrito, o que demonstra que ele não configura necessariamente

⁷⁵ OLIVEIRA, Daniele de Lima. **Deveres e responsabilidade dos administradores da S/A**. Dissertação (Mestrado em Direito) - Faculdade de Direito da Pontifícia Universidade Católica de São Paulo, São Paulo, 2008. p. 57.

⁷⁶ YAZBEK, Otavio. Representação do dever de diligência na doutrina jurídica brasileira: um exercício e alguns desafios. In: KUYVEN, Luiz Fernando Martins (Coord.). **Temas essenciais de direito empresarial: estudos em homenagem a Modesto Carvalhosa**, São Paulo: Saraiva, 2012. E-book.

um novo subdever, mas, sim, pressupõe uma exigência igualmente de monitoramento, pelo administrador, das atividades desenvolvidas na empresa.

3.8 Dever de intervir

Campos define que o dever de intervir ou agir é exigido "quando constatarem algum indício de irregularidade, o que pode se dar não só mediante atuação na operação específica, mas também, sem precipitação, por meio da convocação de reunião da Diretoria, do Conselho de Administração, Assembleia Geral ou Conselho Fiscal."⁷⁷

O dever de intervir, segundo Brigagão⁷⁸, pode ser interpretado como uma decorrência do dever de investigar, o qual estipula que o administrador não pode se manter omissos diante de suas responsabilidades. De fato, o administrador precisa agir de forma proativa e no melhor interesse da companhia sempre com o objetivo de evitar possíveis erros graves.

Contudo, a análise deve ser sempre avaliada com muita cautela, porque poderá haver ocasiões que estrategicamente o administrador decidirá não intervir, como por exemplo quando houver outras questões mais relevantes a demandar a atuação do administrador do que aquela pela qual ele foi acionado, ou, ainda, quando a possível intervenção venha a despender mais energia e custos do que uma eventual não intervenção calculada. Neste caso, todo o contexto que envolveu a possível omissão precisará ser analisado.

⁷⁷ CAMPOS, Luiz Antonio de Sampaio. Deveres e responsabilidades. *In*: LAMY FILHO, Alfredo; PEDREIRA, José Luiz Bulhões (Coord.). **Direito das companhias**. São Paulo: Forense, 2009. p. 1110.

⁷⁸ BRIGAGÃO, Pedro Henrique Castello. **Dever de diligência dos administradores de sociedades anônimas**. Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Direito. Rio de Janeiro, 2012. Disponível em: https://www.puc-rio.br/ensinopesq//relatorio_resumo2012/resumos_pdf.pdf. Acesso em: 17 out. 2022. p. 6.

3.9 Dever de não praticar erros graves

Alguns autores também identificam o dever de não praticar erros graves como subdever do dever de diligência.⁷⁹ Esse subdever não pressupõe a impossibilidade de o administrador cometer qualquer erro, porém preconiza a atuação para evitar erros relevantes, que prejudiquem significativamente a companhia.

Ribeiro define:

Por erros graves de gestão, deve-se entender decisões ou omissões inoportunas ou com risco desproporcional e contrárias à lei, estatuto ou ao interesse social, como a conivência ou omissão com atos ilícitos de outros administradores, inclusive anteriores.

[...]

Constituem também erros relevantes a prática de operações especulativas e arriscadas, com risco elevado ou desproporcional ao benefício esperado, bem como aquelas com falta de garantias ou que impliquem endividamento excessivo para a sociedade.⁸⁰

Tal dever, apesar de ser mencionado por alguns autores, parecer ser, na verdade, um pressuposto para o exercício de qualquer cargo. Ninguém é pago para fazer um trabalho errado. As pessoas são contratadas e pagas pelas empresas para trabalhar corretamente.

Todavia, considerar este subdever como uma decorrência do dever de diligência, permite que, futuramente, o julgador se foque no resultado da tomada de decisão, eis que só após as consequências e resultados é que se saberá se ela foi acertada ou errada.⁸¹ Consoante grande parte da doutrina já discorrida

⁷⁹ Dentre os pesquisados, cita-se exemplificadamente: Renato Ventura Ribeiro, Pedro Henrique Castello Brigagão e Laís Machado Lucas.

⁸⁰ RIBEIRO, Renato Ventura. **Dever de diligência dos administradores de sociedades**. São Paulo: Quartier Latin, 2006. p. 230-231.

⁸¹ Otavio Yazbek entende que os deveres de bem administrar e de não praticar erros graves guardam mais relação com o dever de lealdade do que com o de diligência e atenta para o risco de impossibilidade de aplicação da *business judgment rule* em razão de uma interpretação equivocada do dever de não praticar erros graves.

anteriormente entende, o dever de diligência é um dever de meio, e não de resultado. Então, na medida em que todos os demais subdeveres forem honrados, este dever de não praticar erros graves se esvazia. Isso porque, quando o procedimento adotado para a tomada de decisão foi cauteloso, significa que o administrador foi diligente na tentativa de evitar prejuízos, o que deve ser reconhecido ainda que suas consequências tenham resultado em erro grave para a companhia.

3.10 Dever de participação/assiduidade

Para Brigagão, o dever de assiduidade prescreve que os administradores participem das reuniões de diretoria e do Conselho.⁸² Campos defende como dever de assiduidade: “[...] a obrigação de comparecer às reuniões de Diretoria e de Conselho de Administração [...]”, afirmando que “a ausência reiterada e injustificada [...] denotam falta de diligência por parte dos administradores.”⁸³ Ribeiro e Lucas também enfatizam a existência deste subdever em suas pesquisas.⁸⁴

3.11 Outros deveres envolvidos

Oliveira interpreta que ser diligente é empregar todas “as cautelas, os métodos, recomendações, postulados e diretivas da “ciência” da administração de empresas”⁸⁵, destacando que seu cumprimento exige do “administrador uma postura

YAZBEK, Otavio. Representação do dever de diligência na doutrina jurídica brasileira: um exercício e alguns desafios. In: KUYVEN, Luiz Fernando Martins (Coord.). **Temas essenciais de direito empresarial: estudos em homenagem a Modesto Carvalhosa**, São Paulo: Saraiva, 2012. *E-book*.

⁸² BRIGAGÃO, Pedro Henrique Castello. **Dever de diligência dos administradores de sociedades anônimas**. Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Direito. Rio de Janeiro, 2012. Disponível em: https://www.puc-rio.br/ensinopesq/ccpg/pibic/2012/resumos_pdf/.pdf. Acesso em: 17 out. 2022. p. 6-7.

⁸³ CAMPOS, Luiz Antonio de Sampaio. Deveres e responsabilidades. In: Direito das companhias, coordenado por Alfredo Lamy Filho e José Luiz Bulhoes Pedreira. São Paulo: Ed. Forense, 2009. P. 1110.

⁸⁴ RIBEIRO, Renato Ventura. **Dever de diligência dos administradores de sociedades**. São Paulo: Quartier Latin, 2006. p. 225; LUCAS, Laís Machado. **Programas de integridade nas sociedades anônimas: implementação como conteúdo do dever de diligência dos administradores**. Porto Alegre: Livraria do Advogados, 2021. p. 149.

⁸⁵ OLIVEIRA, Daniele de Lima. **Deveres e responsabilidade dos administradores da S/A**. Dissertação (Mestrado em Direito) - Faculdade de Direito da Pontifícia Universidade Católica de São Paulo, São Paulo, 2008. p. 56.

proativa na fiscalização da companhia, inclusive no que se refere à fiscalização da observância dos princípios da governança corporativa.”⁸⁶ Em seguida, Oliveira defende que os administradores devem permanentemente revisar a estrutura interna da empresa, dando publicidade ao seu sistema de monitoramento para que acionistas e mercado possam validá-lo.

Assim, pode-se interpretar que existem ainda outros três subdeveres: (i) ser proativo na observância dos princípios da governança corporativa, (ii) revisar a estrutura interna da empresa e (iii) dar publicidade ao seu sistema de monitoramento.

Yazbek acrescenta que, a partir da construção jurisprudencial sobre o conteúdo do dever de diligência, surgiu uma crescente exigência e valorização da criação de controles internos adequados e suficientes para o monitoramento e supervisão das atividades da companhia e de seus subordinados, havendo também uma valorização das práticas de *compliance*, como sendo a concretização do dever fiduciário de criar estruturas adequadas. Em sua análise, Yazbek conclui que a “adoção de arranjos institucionais que permitam a verdadeira observância dos deveres de diligência”, chamados por ele de “mecanismos internos de controle” espalharam-se pelo mundo, refletindo na criação de diversos códigos de conduta.⁸⁷

Neste sentido, o pesquisador faz uma avaliação da doutrina brasileira, contextualizando que a implementação de políticas de controle interno também se faz presente no Brasil, o que se pode denotar a partir das recomendações do Instituto Brasileiro de Governança Corporativa (IBGC), que indica a adoção de sistemas de controles internos como forma de monitorar e fiscalizar as atividades empresariais de forma preventiva, prospectiva e proativa.⁸⁸

O autor destaca a importância das obrigações que concretizam o dever de diligência e que auxiliam a sua aplicação na resolução de casos práticos, o que se verifica inclusive em julgamentos da Comissão de Valores Mobiliários (CVM), que, em diferentes oportunidades, reconhecem seus vários aspectos. Todavia, Yazbek

⁸⁶ OLIVEIRA, *loc. cit.*

⁸⁷ YAZBEK, Otávio. Representação do dever de diligência na doutrina jurídica brasileira: um exercício e alguns desafios. In: KUYVEN, Luiz Fernando Martins (Coord.). **Temas essenciais de direito empresarial**: estudos em homenagem a Modesto Carvalhosa, São Paulo: Saraiva, 2012. *E-book*.

⁸⁸ *Ibidem*.

questiona o fato de não haver menção à necessidade de implementação de estruturas de controles internos, nem mesmo quando se avaliam os deveres de fiscalização ou monitoramento. O pesquisador entende que estes controles são relevantes para exame do conteúdo do dever de diligência na atualidade, pois configuram o meio razoável encontrado para a concretização do monitoramento e da fiscalização devidos. O autor refere, inclusive, um novo dever de constituir estruturas adequadas de controle.

Ainda, entende que não reconhecer essa possível atualização do conteúdo do dever de diligência pode relegar à implementação de estruturas de controle a

um papel secundário, como se elas fossem fruto de recomendações ou correspondessem a 'boas práticas de gestão', não ocupando, porém, o núcleo do dever de diligência. Elas seriam, então, formalmente estabelecidas, mas sua capacidade de produzir efeitos no que tange à estrutura da companhia, fora aqueles efeitos mais imediatos e diretamente relacionados ao cumprimento de suas atribuições, seria questionável.⁸⁹

Deste modo, Yazbek sustenta que se deve debater as interpretações contemporâneas do dever de diligência, o que pode vir a produzir efeitos sobre outros dispositivos correlatos. Entende que o reconhecimento da atualização do dever de diligência para incluir em seu conteúdo a necessidade de se atender às recomendações de boa governança é medida que se impõe para lhe assegurar sua efetividade.

No mesmo sentido, Frazão conclui que é dever do administrador adotar normas de *compliance* (prevenção de riscos, canais de comunicação, fluxos de testagem e auditoria da IA), visto pela autora como "o conjunto de ações a serem adotadas no ambiente corporativo para que se reforce a anuência da empresa à legislação vigente, de modo a prevenir a ocorrência de infrações ou, já tendo ocorrido

⁸⁹ YAZBEK, Otavio. Representação do dever de diligência na doutrina jurídica brasileira: um exercício e alguns desafios. In: KUYVEN, Luiz Fernando Martins (Coord.). **Temas essenciais de direito empresarial: estudos em homenagem a Modesto Carvalhosa**, São Paulo: Saraiva, 2012. p. 956. E-book.

o ilícito, propiciar o imediato retorno ao contexto de normalidade e legalidade.”⁹⁰
Frazão destaca que:

o *compliance* reforça a dimensão organizacional do dever de diligência, a fim de que controladores e administradores estruturam a organização empresarial de forma compatível com as atividades da sociedade e com o risco por ela assumido. Para isso, torna-se necessário criar adequados sistemas de vigilância, supervisão e investigação sobre as atividades da sociedade, de modo a assegurar o respeito às obrigações legais e possibilitar a intervenção adequada diante da identificação de problemas e ameaças.

Não é sem razão que um aspecto fundamental dos programas de *compliance* é a necessidade de comprometimento da alta administração. Logo, é inequívoco que o *compliance* acaba ampliando o núcleo básico do dever de diligência, abrindo espaço para que, ao lado do dever de agir bem-informado, acrescente-se igualmente a importante obrigação de instituir e manter uma organização idônea para lidar com o risco assumido, inclusive no que diz respeito à prevenção de ilícitos.⁹¹

Assim, pode-se entender que o *compliance* é o meio pelo qual se colocam em prática as dez condutas diligentes apontadas pela autora.⁹²

Corroborando o acima exposto, a visão de Lucas⁹³ indica que a adoção de programas de integridade não é uma opção para as sociedades anônimas - é uma

⁹⁰ FRAZÃO, Ana. Responsabilidade civil dos administradores de sociedades empresárias por decisões tomadas com base em sistemas de inteligência artificial. In: FRAZÃO, Ana; MULHOLLAND, Caitlin (Coord.). **Inteligência artificial e Direito: ética, regulação e responsabilidade**. 2. ed. rev. atual. e ampl., p. 481-525, São Paulo: Thomson Reuters Brasil, 2019. p. 523.

⁹¹ FRAZÃO, Ana. Responsabilidade civil dos administradores de sociedades empresárias por decisões tomadas com base em sistemas de inteligência artificial. In: FRAZÃO, Ana; MULHOLLAND, Caitlin (Coord.). **Inteligência artificial e Direito: ética, regulação e responsabilidade**. 2. ed. rev. atual. e ampl., p. 481-525, São Paulo: Thomson Reuters Brasil, 2019. p. 523-524.

⁹² Conforme já relatado em tópico anterior, são elas: (i) dever de informação para a tomada de decisões, (ii) competência profissional para o cargo, (iii) perícia compatível com as funções desempenhadas, (iv) conhecimento sobre os negócios da companhia, (v) estruturação de uma organização empresarial compatível com o porte e risco assumido pela cia, (vi) estruturação de um sistema de vigilância, supervisão e investigação, (vii) dever de intervir diante de problemas ou ameaças, (viii) respeito às obrigações legais, sem descumprimentos “calculados”, (ix), presença nas reuniões dos colegiados que é parte e (x) utilização de processos decisórios justos, adequados e razoáveis. *Ibidem*.

⁹³ LUCAS, Laís Machado. **Programas de integridade nas sociedades anônimas: implementação como conteúdo do dever de diligência dos administradores**. Porto Alegre: Livraria do Advogados, 2021.

obrigatoriedade decorrente do dever de diligência dos administradores, conforme previsto no art. 153 da Lei 6.404/76.

Apesar de fazer sua análise com foco nos programas incentivados pela Lei Anticorrupção, a autora ressalva que “o conteúdo do dever de diligência dos administradores é mutante, podendo ser alterado, sempre que surjam novas legislações ou práticas de mercado que exijam a revisão ou readaptação da conduta dos administradores”⁹⁴, conduta esta que deve ser adotada com a intenção de (i) evitar prejuízos financeiros advindos de multas, de condenações, entre outros, (ii) evitar exposição pública e midiática indesejável, em razão de alguma prática incorreta da empresa e (iii) evitar riscos desmedidos.

Ao investigar profundamente o conteúdo do dever de diligência dos administradores de sociedades anônimas, Lucas destaca que:

o dever de diligência remete ao cuidado, zelo e empenho que o homem ativo e probo empregaria na condução dos seus próprios negócios. A implementação dos programas de integridade materializa essas condutas, sob vários aspectos. [...] um primeiro aspecto a ser analisado é que os programas de integridade atuam como mitigadores de riscos das empresas. [...] um segundo aspecto que coloca a implementação dos programas de integridade dentro do escopo do dever de diligência é o auxílio que esses podem dar para que o administrador cumpra as atuais exigências do dever de fiscalizar. [...] o terceiro aspecto guarda relação com a proteção e as expectativas do mercado de capitais em relação à companhia e as regras impostas pelo seu órgão regulador, a Comissão de Valores Mobiliários (CVM).⁹⁵

A autora também reforça que “a implementação dos programas de integridade pode gerar o efeito reflexo da isenção ou da diminuição da responsabilidade dos administradores por terem eles empreendido todos os esforços para o efetivo monitoramento da empresa.”⁹⁶

⁹⁴ *Ibidem*, p. 153.

⁹⁵ LUCAS, Laís Machado. **Programas de integridade nas sociedades anônimas**: implementação como conteúdo do dever de diligência dos administradores. Porto Alegre: Livraria do Advogados, 2021. p. 153-155; 157.

⁹⁶ *Ibidem*, p. 159.

3.12 Business Judgment Rule

Internalizada no ordenamento brasileiro a partir do art. 159, § 6º da Lei das SA⁹⁷, a *Business Judgment Rule* está positivada da seguinte maneira: “o juiz poderá reconhecer a exclusão da responsabilidade do administrador, se convencido de que este agiu de boa-fé e visando ao interesse da companhia”. Segundo Mariana Pargendler, a *Business Judgment Rule* é um antídoto contra a discricionariedade decisória para responsabilização dos administradores, eis que impõe uma presunção de que estes agiram de maneira informada, de boa-fé e na crença de que atendiam os melhores interesses para a companhia, atendendo aos deveres de diligência e lealdade.⁹⁸

A *Business Judgment Rule* é o foco de análise de Campos, que explica:

O administrador é livre para decidir sobre a conveniência e oportunidade dos negócios sociais, podendo, naturalmente, ter sucesso ou não na sua decisão. De seu eventual erro ou insucesso na tomada de decisão não repercutirá qualquer responsabilidade, porquanto tenha decidido de maneira diligente, em boa-fé e naquilo que supôs ser o melhor interesse social. É que o erro de gestão, por si só, em regra, não traz responsabilidade para o administrador.

[...]

A verificação do cumprimento ou não do dever de diligência reside mesmo em uma avaliação de procedimento, e não do conteúdo da decisão.⁹⁹

Konrad e Lupion mencionam que os administradores estão sujeitos aos deveres e responsabilidades previstos na Lei 6.404/76 e que “A decisão empresarial, regularmente tomada, vincula a companhia e não atinge os administradores, ainda

⁹⁷ Há autores como Brigagão, que entendem que a previsão da *Business Judgment Rule* encontra sua previsão legal no art. 158. BRIGAGÃO, Pedro Henrique Castello. **A administração das companhias e a business judgment rule: uma análise à luz do direito brasileiro.** 191 p. Monografia (Graduação em Direito). Pontifícia Universidade Católica do Rio de Janeiro: Rio de Janeiro, 2014, p. 130.

⁹⁸ PARGENDLER, Mariana. Responsabilidade civil dos administradores e business judgment rule no Direito Brasileiro. **Revista dos Tribunais**, São Paulo, v. 953, p. 51-74, 2015.

⁹⁹ CAMPOS, Luiz Antonio de Sampaio. Deveres e responsabilidades. In: LAMY FILHO, Alfredo; PEDREIRA, José Luiz Bulhões (Coord.). **Direito das companhias.** São Paulo: Forense, 2009. p. 1101-1105.

que tenha causado prejuízos e impactado negativamente a empresa”¹⁰⁰, eis que se enquadra na hipótese criada para adaptação do direito norte-americano (*Business Judgment Rule*) para a realidade brasileira.

Os autores referem que a *Business Judgment Rule* ou regra do julgamento do negócio representa uma segurança para os conselheiros e diretores que tomam decisões fundamentadas em informações confiáveis, os quais devem fazer constar em ata que buscaram os subsídios adequados e suficientes para a tomada de decisão. Destacam que os próprios administradores é que devem buscar mais dados para obter informações relevantes e disponíveis. Também alertam que:

A complexidade e o profissionalismo da atividade empresarial, que concentram na sociedade anônima seu maior exemplo, demonstram que o Conselho de Administração e a Diretoria devem criar estruturas e mecanismos para o funcionamento da companhia. Pode-se dizer que os administradores estarão em “portos seguros” se atentarem para a conjugação do dever de diligência e da regra da decisão empresarial quando tiverem que decidir negocialmente sobre determinado contrato empresarial, já que foco, no exame da *Business Judgment Rule*, é se houve indícios de má-fé, fraude, interesse ou conflito do administrador ou falha de procedimento no processo decisório, já que as suas decisões devem ser consideradas regulares porque gozam da presunção de terem sido tomadas no melhor interesse da companhia.¹⁰¹

Ao relacionar o dever de diligência com a *Business Judgment Rule*, Silva ressalta que, se a tomada de decisão do administrador foi devidamente informada, não há espaço para revisão judicial, mesmo que sua ação acarrete perdas para empresa. Neste sentido, refere que caberá ao contestante comprovar que os administradores atuaram inadvertidamente ou com desconhecimento indesculpável. Esclarece:

¹⁰⁰ KONRAD, Carlos Bender; LUPION, Ricardo. Dever de diligência dos administradores em tempos de pandemia e impactos da crise nos contratos empresariais. *RJLA*, [S.l.], ano 7, n. 3, p. 427-456, 2021. p. 440.

¹⁰¹ *Ibidem*, p. 443.

O dever de diligência foca no processo de tomada de decisão com análise de possibilidades e alternativas, bem como na identificação da matéria em questão. Entretanto, não há fórmula preestabelecida ou padrão que o Conselho de Administração deva seguir para tomar uma decisão e cumprir o dever de diligência. Os riscos (e custos) associados para se obter informação devem ser sopesados contra os benefícios. [...] A quantidade de informação necessária ou requerida para a tomada de decisão também é um julgamento ou decisão empresarial, o que faz esta decisão sujeita à *Business Judgment Rule*.¹⁰²

Maria Eduarda Ferreira Piccoli e Oksandro Osdival Gonçalves examinaram os aspectos mencionados acima, citando a observância do dever de diligência como um dos elementos para manutenção da presunção da *Business Judgment Rule*, destacando que o “descumprimento do dever legal específico gera a ilicitude e, na linha exposta anteriormente, subsidia, sem a possibilidade de aplicação da *Business Judgment Rule*, a responsabilidade civil.”¹⁰³

Os autores revelam, ainda, que a violação ao dever de diligência dos administradores de companhias abertas é a acusação mais frequente em julgamentos da Comissão de Valores Mobiliários (CVM), que acabam por absolver a grande maioria dos administradores¹⁰⁴, reconhecendo que o insucesso dos resultados esperados é inerente aos riscos da atividade empresarial, e que o cumprimento do dever de diligência se consolida no transcurso do processo decisório.¹⁰⁵

Por outro lado, referem que “caso o administrador provoque prejuízos com culpa ou dolo ou promova qualquer infração da lei ou do estatuto social restará caracterizada a irregularidade do ato de gestão e, portanto, haverá responsabilidade pessoal do administrador.”¹⁰⁶

¹⁰² SILVA, Alexandre Couto. **Responsabilidade dos administradores de S/A: business judgment rule.** Rio de Janeiro: Elsevier, 2017. p. 200.

¹⁰³ PICCOLI, Maria Eduarda Ferreira; GONÇALVES, Oksandro Osdival. A responsabilidade civil dos administradores no Brasil e a Business Judgment Rule. **Revista Jurídica Luso-Brasileira - RJLB, [S.l.]**, ano 8, n. 6, p. 1497-1530, 2022. p. 1517.

¹⁰⁴ A Comissão de Valores Mobiliários, em decisões unânimes, reluta, na maior parte dos casos, em rever e reverter decisões da administração, quando demonstrada a diligência na sua construção. *Ibidem*, p. 1522.

¹⁰⁵ *Ibidem*, p. 1519-1520.

¹⁰⁶ *Ibidem*, p. 1514.

Como se vê, a *Business Judgment Rule* reforça a importância de o administrador cumprir com o dever de diligência. E mais: de se comprovar o cumprimento do dever de diligência, que, juntamente com a comprovação da boa-fé e da lealdade, garantem o respeito à decisão negocial. Neste sentido, adiante avaliar-se-á, na prática, a aplicação deste direito e como ele impacta na avaliação do dever de diligência, a partir de julgados da CVM.

3.13 A Comissão de Valores Mobiliários (CVM) e o dever de diligência

Como se viu acima, o dever de diligência é objeto de muitas interpretações, o que dificulta compreender objetivamente quais são as condutas que resguardam um administrador que desejar ser considerado diligente perante um caso concreto. Deste modo, buscou-se examinar como a Comissão de Valores Mobiliários julgou situações concretas.

Para esta análise, pesquisou-se no sítio eletrônico da CVM¹⁰⁷, onde se constatou a existência de mais de 102 julgamentos relativos a 98 processos sancionadores envolvendo os termos "art. 153", "artigo 153", "dever de diligência", "dever de diligencia", "diligência" e "diligencia" nos últimos cinco anos.

Considerando o grande volume de resultados, houve a necessidade de fazer uma análise por amostragem, optando-se por se debruçar sobre três casos emblemáticos que envolveram extensos debates e que receberam grande notoriedade, como os Processos Administrativos Sancionadores (PAS) de n. 19957.010647/2019-97 (05/2016 - assunto: Refinaria Abreu e Lima – RNEST), n. 19957.011654/2019-14 (06/2016 - assunto: Complexo Petroquímico do Rio de Janeiro – Comperj) e n. 19957.005789/2017-71 (SP2017/294 - assunto: *impairment* na elaboração das demonstrações financeiras dos exercícios sociais encerrados em 31/12/2010, 31/12/2011, 31/12/2012, 31/12/2013 e 31/12/2014), envolvendo a empresa Petrobrás.¹⁰⁸

¹⁰⁷ Pesquisa realizada em: 30/07/22. PROCESSOS sancionadores julgados. In: COMISSÃO DE VALORES MOBILIÁRIOS. Brasília, 2022. Disponível em: <https://conteudo.cvm.gov.br/sancionadores>. Acesso em: 17 out. 2022.

¹⁰⁸ Todos os votos se encontram consolidados e disponíveis em: CVM conclui julgamento de processos envolvendo Petrobras. In: COMISSÃO DE VALORES MOBILIÁRIOS. Brasília, 03 nov. 2020.

Após transcrever o relatório fático dos julgamentos nas palavras do relator dos casos, o diretor Henrique Machado, a pesquisa cuidará exclusivamente de destacar os pontos considerados de maior pertinência para interpretação do conteúdo do dever de diligência, sem adentrar na análise de outros temas debatidos nos julgamentos (como dever de lealdade, prescrição, entre outros):

1) 19957.010647/2019-97 (05/2016 - assunto: Refinaria Abreu e Lima – RNEST):

A RNEST surgiu, em 14.02.2005, como projeto conjunto da Petrobras e da companhia de petróleo venezuelana PDVSA, a partir da assinatura de um memorando de entendimentos para o desenvolvimento de negócios e atividades de cooperação no setor de refino, que tinha por objetivo atender a estratégia de processamento de petróleo do Brasil e da Venezuela (fls. 477-479). (...) Naquele primeiro momento, a Petrobras previu que o valor total do investimento seria de US\$2,3 bilhões, com VPL de US\$209,9 milhões. (...) 102. Pouco tempo depois, em 08.03.2007, a diretoria da Petrobras aprovou a implementação do Plano de Antecipação da Refinaria pela qual tomava medidas para antecipar o início das atividades da RNEST para o mês de agosto de 2010. Houve, a partir desse momento, a necessidade de acelerar as diversas contratações que somente deveriam ocorrer na fase de execução do projeto.

103. Após a fase de planejamento, em 25.11.2009, a diretoria autorizou o início da fase de execução do projeto, nos termos do documento interno DIP 327/2009, com a previsão de custo modificada de US\$4,1 bilhões para US\$13,3 bilhões. Anteriormente a essa aprovação, o DIP 212/2009, de 3.9.2009, trazia para o empreendimento, um VPL negativo em US\$3,067 bilhões, o que motivou a alteração de determinadas premissas, consignadas no DIP 327/2009, que elevaram o VPL para um valor positivo de US\$76 milhões, quando o projeto teve sua execução aprovada pela diretoria. (...) Como bem detalhado no relatório, a Acusação questiona a diretoria por ter aprovado, em 08.03.2007, a antecipação das contratações para a construção da Refinaria. Segundo afirma, os diretores teriam ignorado a Sistemática Corporativa e os alertas das áreas técnicas quanto à falta de definição da parceria com a PDVSA e à reduzida margem de segurança

do cronograma de construção da Refinaria, para decidir o início das atividades operacionais da RNEST. (...) Diante disso, a acusação concluiu que o conselho de administração não teria discutido a matéria à luz dos sinais de alerta relativos à inviabilidade da RNEST, ratificando a decisão da diretoria sem maiores discussões, o que caracterizaria violação ao dever de diligência previsto no art. 153 da Lei nº 6.404/76.

2) 19957.011654/2019-14 (06/2016 - assunto: Complexo Petroquímico do Rio de Janeiro – Comperj):

Conforme relatado, este processo foi instaurado pela SPS para apurar a responsabilidade dos administradores da Petrobras por supostamente descumprirem seus deveres fiduciários nos procedimentos relativos à construção do COMPERJ. (...) O COMPERJ surgiu em 2004 a partir da ideia de implantar um complexo destinado à produção de petroquímicos derivados de petróleo pesado oriundo da Bacia de Campos. Inicialmente, o complexo seria formado por uma unidade de refino básica destinada à produção de petroquímicos básicos, os quais seriam utilizados como insumos das unidades petroquímicas encarregadas de produzir resinas e plásticos. Diante da complexidade e grandeza do empreendimento, foi prevista a construção de uma central de utilidades para suprir o complexo com água, vapor, combustíveis, energia elétrica, entre outros insumos, assim como instalações prediais para viabilizar o transporte e a estocagem de matérias-primas e produtos. As decisões sobre projetos de investimentos como o COMPERJ seguiam, à época, norma metodológica de gestão denominada Sistemática Corporativa. De acordo com esse normativo aprovado pela diretoria executiva, a condução dos projetos de investimentos é dividida em cinco fases, conforme o nível de especificação e maturidade do projeto. Ao final de cada uma das fases de planejamento, existem três marcos de avaliação, com pré-requisitos distintos, para prosseguir à fase seguinte. Esses marcos são definidos a fim de que os administradores possam avaliar a pertinência de prosseguir, cancelar, adiar ou reformular o projeto, antes do emprego de recursos financeiros, evitando custos não programados decorrentes de alterações no projeto. (...) Em 6.9.2006, a diretoria da Petrobras autorizou que o COMPERJ passasse da fase II para a fase III e que gastos para aquisição de equipamentos fossem antecipados. Em 26.02.2010, a diretoria aprovou a alteração do modelo conceitual do COMPERJ, que passou a ser gerido na forma de "programa" composto por três etapas, com projetos distintos e

interdependentes entre si. Dentro do programa, também foi autorizado o começo da fase IV da UPB da Refinaria. Segundo a Acusação, os administradores teriam faltado com os seus deveres fiduciários por ocasião dessas duas deliberações. [...]

3) 19957.005789/2017-71 (SP2017/294 - assunto: *impairment* na elaboração das demonstrações financeiras dos exercícios sociais encerrados em 31/12/2010, 31/12/2011, 31/12/2012, 31/12/2013 e 31/12/2014),

Trata-se de PAS instaurado pela SEP para apurar a responsabilidade de ex-membros da Diretoria, do Conselho de Administração e do Conselho Fiscal da Petrobras por infração a dispositivos da Lei nº 6.404/1976 e da Instrução CVM nº 480/2009, em razão do suposto descumprimento, pela administração da Companhia, quando da elaboração das demonstrações financeiras de 31.12.2010, 31.12.2011, 31.12.2012, 31.12.2013 e 31.12.2014, de norma contábil que dispõe sobre procedimentos de verificação e reconhecimento da redução ao valor recuperável de ativos. (...) A discussão de mérito travada nos autos versa, basicamente, sobre o cumprimento ou não dos requerimentos do CPC 01 (R1) nos testes de recuperabilidade – ou testes de *impairment* – de ativos de refino da Petrobras, realizados para as demonstrações financeiras dos exercícios sociais de 2010 a 2014, e a apuração das responsabilidades por eventuais infrações à Lei nº 6.404/1976 e à Instrução CVM nº 480/2009, em virtude de algum descumprimento do pronunciamento técnico. O objetivo do CPC 01 (R1), explicitado em seu item 1, é estabelecer procedimentos a serem seguidos pelas entidades a fim de assegurar que seus ativos estejam registrados contabilmente por valores que não excedam os valores de recuperação.

Como já mencionado, os três julgamentos geraram intensos debates entre os diretores, tanto que as decisões foram definidas por maioria. Apesar de reforçarem os entendimentos teóricos já abordados nesta pesquisa, os julgamentos permitem algumas reflexões.

O relator dos casos entendeu haver descumprimento do dever de diligência da diretoria da Petrobrás, na medida em que aprovou antecipações de projetos que possuíam possibilidade evidente de prejuízo, mesmo com todas as informações e alertas demonstrados por auditorias e áreas especializadas. Em semelhante sentido, o relator compreendeu que o Conselho de Administração não exerceu o dever de

fiscalização ao desconsiderar os sinais de alerta apresentados em pareceres técnicos e não indagar a diretoria sobre a decisão tomada.

Um dos pontos que merece destaque é a ressalva feita, pelo relator, a respeito da ausência de registro formal de debates demonstrando que as decisões foram tomadas de forma refletida, bem como de justificativas que embasassem a tomada de decisão que assumiu grandes riscos, apesar de todos os sinais de alerta disponíveis.

Por sua vez, os diretores Gustavo Machado Gonzales e Flavia Perlingeiro compreenderam que a falta de registros ou justificativas formalizadas não presume a ausência de uma decisão informada e refletida, tendo em vista que todos os relatórios e pareceres disponibilizados aos diretores e conselheiros não comprovam que não houve debates (apenas não detalham eles). Relatam que, na época dos fatos, não era uma praxe empresarial a descrição minuciosa de todos os temas discutidos e decididos em reuniões de diretoria e do conselho.

Deste modo, não havendo provas de má-fé ou de ausência de reflexão a respeito das informações disponíveis, entenderam que não lhes cabia avaliar *a posteriori* as decisões tomadas, respeitando-se a discricionariedade que deve ser resguardada aos administradores, de forma a não desestimular a assunção de riscos – em respeito a *Business Judgment Rule*. Assim, se posicionaram em favor do direito de os administradores tomarem decisões arriscadas, apesar de todos os potenciais e conscientes danos de alguma decisão.

Em que pese os julgamentos tenham concluído que o dever de diligência não foi violado, alguns elementos dão a entender que, se a situação avaliada ocorresse atualmente, outro desfecho poderia ter ocorrido, em virtude da alteração das práticas de mercado (como o maior desenvolvimento do *compliance* no Brasil, especialmente após a operação Lava-Jato):

Creio que nessas passagens transparecem três graves erros da Acusação, que explicam suas conclusões, mas que tornam impositiva a absolvição dos acusados. O primeiro é tomar o dever de monitoramento como uma obrigação de resultados, quando é pacífico tratar-se de obrigação de meios. O segundo é ignorar que a diligência deve ser avaliada à luz da época dos fatos – e, nesse

ponto, parece-me forçoso reconhecer que embora o compliance não seja novidade no Brasil, ele se desenvolveu de modo formidável sobretudo a partir (e, em larga medida, em razão) dos graves episódios descortinados pela Operação Lava-Jato. O terceiro problema, na minha avaliação, é que a Acusação se baseia em expectativas bastante irreais, como se controles internos tivessem eficácia imediata e efetividade absoluta, quando se sabe que o processo de adoção de controles internos é complexo e demorado, demandando uma série de medidas, como, por exemplo, o treinamento dos funcionários, e que sempre estará sujeito a falhas e dependente de melhorias.¹⁰⁹

Isto reforça o que se vê a partir do exame doutrinário anteriormente dissertado, de que as práticas que concretizam o cumprimento do dever de diligência devem ser delineadas à luz do caso concreto e do contexto em que a empresa, o administrador e o negócio decidido estão inseridos, sob pena de seguir um “script” que se esvazia diante de uma circunstância atípica ou inovadora, como pode ocorrer no uso de inteligência artificial.

3.14 Conclusões parciais

Constata-se a partir da doutrina analisada de maneira cronológica que a interpretação do que significa cumprir o dever de diligência se adaptou conforme (i) o passar do tempo, (ii) a evolução das legislações, (iii) as práticas do mercado, (iv) as informações que se tornaram disponíveis a partir do aprimoramento técnico, do desenvolvimento de novas profissões e especialidades, entre outras variáveis.

Tal circunstância só foi permitida, pois, o dever de diligência tem previsão legal com conceitos em aberto, cujo conteúdo é propositalmente elástico para abarcar novas condutas que se tornem cogentes para os administradores conforme o contexto de cada empresa, analisando-se tempo, local e recursos da tomada de decisão.

¹⁰⁹ BRASIL. Comissão de Valores Mobiliários. **Processo Administrativo Sancionador CVM nº 05/2016**. Processo Eletrônico SEI nº 19957.010647/2019-97. Voto Diretor Gustavo Machado Gonzales. Fls. 45-46. 2020. Disponível em: <https://www.gov.br-br/assuntos/noticias.pdf>. Acesso em: 17 out. 2022.

A subjetividade da previsão legal da Lei das SA permite que em cada caso concreto se analise seu cumprimento, não havendo um “*checklist*” fixo de padrões pré-estabelecidos para todos os casos.

Ademais, também se verificou que seu cumprimento é uma obrigação de meio e não de resultado, cabendo ao intérprete analisar se cautelas foram tomadas antes e durante uma tomada de decisão com potencial de risco.

Ainda, foi possível denotar algumas divergências doutrinárias, tanto com relação ao modelo a ser seguido como padrão para cumprimento do dever de diligência – que alguns apontam como sendo o bom pai de família enquanto outros apontam como sendo o homem de negócios –, quanto com relação aos seus subdeveres (dever de se informar, de vigiar, de monitorar etc.), cujo reconhecimento varia conforme a interpretação de cada autor.

O que pode ser afirmado é que as condutas esperadas por um administrador diligente são muitas. Nas pesquisas mais atuais, é constante a menção a um dever de *compliance*, de boas práticas de governança corporativa, tendo em vista que elas concretizam e formalizam muitos dos subdeveres.

A prática de formalizar, ou seja, de registrar e de documentar todas as etapas e as evidências que embasam um processo decisório também pode ser entendida como um subdever, o que se faz a partir do estudo de Konrad e Lupion e do voto do relator dos casos da Petrobrás (apesar de vencido, os administradores correram sérios riscos de serem condenados por não terem provas documentais da diligência empregada antes da tomada de decisão que prejudicou a saúde financeira da empresa).

O administrador tem que agir como se suas ações/decisões fossem ser auditadas futuramente, e por isso deve produzir os meios e justificativas que lhes resguardem. Assim, por mais que futuramente a decisão se torne danosa para a companhia, haverá justificativas para comprovar que a tomada de risco foi feita de maneira diligente, forte em premissas e critérios válidos.

Portanto, quando se pensar em uma estrutura de *compliance* empresarial, o registro de todos os passos tomados e, principalmente, das justificativas para as ações e não ações devem ser documentadas. Seguir a premissa de “não basta ser

[diligente], tem que parecer [diligente]" é o norte do administrador que queira ser reconhecido como cumpridor do dever de diligência.

Deste modo, retomando a pergunta que centrou este capítulo: no que consiste o dever de diligência?

Entende-se que para uma tomada de decisão diligente, ou seja, para um fechamento de novo contrato de qualquer natureza, adoção de novas práticas e ferramentas no ambiente corporativo (como o uso de IA, por exemplo), definição de custos e investimentos etc., o administrador deve previamente: reunir informações e documentos sobre o assunto envolvido na decisão a ser tomada. Estas informações e documentos precisam ser qualificados, ou seja, oriundos de fontes seguras e confiáveis (mídia reconhecidamente confiável, pesquisas acadêmicas, livros, documentários). Depois, é necessário investigá-los, ou seja, examinar criteriosamente as informações obtidas, buscando compreender aspectos legais, éticos, práticos, financeiros, sociais etc., do assunto envolvido, devendo contar com apoio de especialistas quando não dispor de conhecimento técnico para tanto. Se, nesta análise, o administrador verificar a existência de potenciais danos, deve avaliá-los com mais profundidade, verificando alternativas e soluções para contornar os obstáculos ou, então, justificar plausivelmente porque o risco é válido para a empresa (desde que não acarrete descumprimentos legais ou danos intencionais).

Além disso, mesmo após a decisão tomada, o administrador deve monitorar o seu objeto constantemente. Ou seja, supervisionar (diretamente ou por intermédio de terceiros) o andamento do contrato/atividade implementada para acompanhar as consequências e os resultados, recalculando rotas e adotando melhorias em tempo hábil sempre que se perceber a necessidade de algum ajuste na decisão tomada. E sem esquecer de manter registros e evidências idôneos de todas estas etapas, quase como um prontuário. Tudo isso deve estar no foco de um programa de *compliance*, sendo sua estrutura e dimensão adequadas às peculiaridades e recursos disponíveis para cada empresa.

Feitas estas considerações, no próximo capítulo será analisado como essa diligência deve ser observada pelas empresas que desenvolvem ou utilizam inteligência artificial em alguma etapa de suas atividades, de modo a evitar ou mitigar danos decorrentes de discriminação algorítmica.

4. A DISCRIMINAÇÃO ALGORÍTMICA E O DEVER DE DILIGÊNCIA

Como visto no decorrer do presente trabalho, o uso de inteligência artificial na tomada de decisões e no desempenho de atividades empresariais pode ocasionar diversos efeitos colaterais, sendo o efeito objeto do presente estudo a discriminação algorítmica, cujo impacto na vida das pessoas afetadas pode ser muito danoso por ocasionar restrição de serviços e até de direitos.

Tais efeitos eram desconhecidos há alguns anos, tendo em vista a expansão recente do uso de IA e, conseqüentemente, do acompanhamento e estudo de seus impactos, de modo que muitos administradores de empresas não puderam tomar as devidas cautelas previamente ao seu uso por pura imprevisibilidade.

Contudo, esta não é a situação atual. Muitos são os casos que despertam o alerta para o uso de IA: como visto no primeiro capítulo, se o seu ciclo de vida não for bem arquitetado, pode se tornar um verdadeiro "algoritmo de destruição em massa", valendo-se das palavras de O'Neil.

Assim, a esta altura do estado da arte, é seguro afirmar que a adoção de algoritmos de inteligência artificial pelas empresas pode ser uma situação potencialmente danosa quando não forem adotadas cautelas mínimas desde o seu planejamento e sua projeção.

Conjugando esta constatação com o que foi estudado a respeito do dever de diligência, é possível concluir que este dever precisa ser observado pelos administradores antes, durante e após a adoção de processos decisórios que envolvam inteligência artificial.

Deste modo, o administrador deve cumprir os subdeveres de diligência desde a escolha da atividade que será executada, apoiada ou substituída por IA, até a escolha (i) dos responsáveis por sua implementação e desenvolvimento, (ii) das regras decisórias que orientarão o "agir" algorítmico, (iii) dos dados que serão utilizados para treinar a máquina, entre outras etapas necessárias, até a fase de testagem, auditoria, monitoramento, ajustes e aperfeiçoamento.

Sendo assim, todas estas decisões que acabam sendo tomadas antes e durante a implementação de uma ferramenta de IA assumem o caráter de uma

decisão negocial, haja vista que seu objetivo é sempre o de trazer benefícios (como o atendimento a um maior número de clientes) ou reduzir custos para as empresas (decorrentes, por exemplo, da substituição de colaboradores por RPAs¹).

Nesta medida, tais decisões negociais, tomadas pelos administradores das empresas direta ou indiretamente (por meio de prepostos humanos ou algorítmicos), estão sujeitas ao cumprimento do dever de diligência. Frazão compreende que se trata de uma delegação da atividade empresarial a um sistema de inteligência artificial, sob a justificativa de se obter uma decisão mais assertiva.²

Maria Cristina Borrasca Felisberto também analisa o dever de diligência empregado no uso de inteligência artificial, destacando que o limite de atuar conforme o dever de diligência não seria observado nas decisões tomadas pela IA em si, mas na tomada de decisão de qual(is) atividade(s) devem ou podem ser delegadas ao sistema (e sob que circunstâncias e cuidados), tendo em vista os riscos e responsabilidades que qualquer delegação atrai.³

A autora destaca que os principais pontos de observação são os dados coletados e o processamento destes dados, cabendo ao administrador se acautelar com relação à qualidade dos dados que abastecem o sistema, sua origem, veracidade, exatidão, precisão, acurácia e destinação, robustez, inteligibilidade, objetividade, imparcialidade, neutralidade.⁴

Com relação ao processamento dos dados, em razão de sua complexidade e opacidade, difícil de ser interpretada até por desenvolvedores e especialistas no tema, a autora compreende que não se exige do administrador um conhecimento tão aprofundado, eis que isto extrapolaria o próprio conceito de diligência trazido pelo art. 153 da Lei nº 6.404/76⁵ (salvo se a natureza do próprio negócio for o desenvolvimento de sistemas de IA). Contudo, compreende que

¹ Sigla para o termo *Robotic Process Automation* (automação de processos robóticos).

² FRAZÃO, Ana. Responsabilidade civil dos administradores de sociedades empresárias por decisões tomadas com base em sistemas de inteligência artificial. In: FRAZÃO, Ana; MULHOLLAND, Caitlin (Coord.). **Inteligência artificial e Direito: ética, regulação e responsabilidade**. 2. ed. rev. atual. e ampl., p. 481-525, São Paulo: Thomson Reuters Brasil, 2019. p. 484-485.

³ FELISBERTO, Maria Cristina Borrasca. Business judgment rule: o limite dos riscos assumidos pelos administradores e o uso da Inteligência Artificial. **Revista Brasileira de Direito Comercial**, [S.l.], n. 43, out./nov. 2021.

⁴ *Ibidem*, p. 102.

⁵ *Ibidem*, p. 106.

não se dispensa o conhecimento da tecnologia a ser empregada, os algoritmos que serão utilizados, e, até mesmo, a forma como eles irão se relacionar com os dados a fim de se obter o resultado pretendido. Além disso, deve o administrador, ante as informações que obtiver, se antever a eventuais riscos que ele pode gerar, e buscar formas de preveni-los. E tais precauções aplicam-se tanto a sistemas desenvolvidos internamente, como adquiridos no mercado.⁶

Cada vez mais os administradores são incentivados a utilizar IA ou como apoio a seus diagnósticos e análises, ou até como substituição de determinadas atividades empresariais. Contudo, como visto, os algoritmos podem ser falíveis e isto envolve risco:

[...] mesmo quando o raciocínio estatístico estiver certo, é forçoso concordar em que este, em si, não deixa de representar uma espécie de discriminação quando utilizado em relação a pessoas, na medida em que o julgamento a respeito de determinado sujeito é feito a partir de critérios gerais que desconsideram a sua individualidade. Por essa razão, em que pese a utilidade e a importância da estatística, há que se tomar cuidado com o seu emprego amplo e absoluto, sem que existam mecanismos ou válvulas de escape para a correção dos resultados.⁷

Neste sentido, Frazão alerta que deve haver parcimônia e prudência com tais sistemas em todos os aspectos da vida, inclusive nas atividades empresariais. Refere que muitas empresas adotam tais ferramentas sem a devida transparência e *accountability*⁸, considerando que sua ferramenta pode estar protegida por segredo

⁶ *Ibidem*, p. 104.

⁷ FRAZÃO, Ana. Responsabilidade civil dos administradores de sociedades empresárias por decisões tomadas com base em sistemas de inteligência artificial. In: FRAZÃO, Ana; MULHOLLAND, Caitlin (Coord.). **Inteligência artificial e Direito: ética, regulação e responsabilidade**. 2. ed. rev. atual. e ampl., p. 481-525, São Paulo: Thomson Reuters Brasil, 2019. p. 486.

⁸ Vale-se das palavras de Andriei Gutierrez para definir o conceito de *accountability*: Sem tradução exata para a língua portuguesa, trata-se de um conceito do idioma inglês que abarca práticas que remetem à responsabilidade com ética, à obrigação, à busca por transparência, à prestação de contas. De maneira simplificada, significa que aqueles que desempenham funções relevantes na sociedade deveriam dar transparência ao que estão fazendo, por quais motivos, e como estão fazendo. Remete à necessidade de governança e, em alguns casos, até responsabilidade civil. GUTIERREZ, Andriei. É possível confiar em um sistema de Inteligência Artificial? Práticas em torno da melhoria da sua confiança, segurança e evidências de *accountability*. In: FRAZÃO, Ana. Responsabilidade civil dos administradores de sociedades empresárias por

industrial. Contudo, mesmo diante de tais proteções, a autora recomenda que haja algum tipo de controle sobre a qualidade dos dados utilizados para treinar o algoritmo, analisando sua veracidade, exatidão, precisão, acurácia, adequação e pertinência para os fins destinados, bem como a qualidade no processamento destes dados, assegurando que a programação é idônea e apta a produzir resultados confiáveis.

Apesar de tais necessidades, verifica-se que muitas empresas têm adotado o uso de IA sem atentar para as devidas cautelas. Todavia, não se pode ignorar que isto pode configurar uma temeridade, eis que:

[...] imputa-se ao administrador as decisões dos sistemas de inteligência artificial, mesmo em casos de completa automação, uma vez que foi o administrador que, ao cancelar o resultado dos referidos sistemas, encampou aquela decisão como se sua fosse. [...] Logo, quando agem de forma ilícita, os administradores automaticamente deflagram, ao lado da sua responsabilidade pessoal, a responsabilidade da pessoa jurídica. [...] A introdução da inteligência artificial, como já se adiantou, em nada altera o tratamento da matéria, já que, em qualquer circunstância de delegação de suas competências para sistemas de inteligência artificial, as decisões ou atos serão imputados ao administrador e o sujeitarão à responsabilidade pessoal, nos termos dos critérios já previstos em lei.⁹

Frazão também menciona que tal responsabilidade dos administradores é fundada na culpa, devendo-se avaliar se o administrador descumpriu algum dever jurídico que estava obrigado, como, por exemplo, o dever de diligência.

Assim, a autora entende que, quando a empresa empregar o uso de inteligência artificial, o administrador pode consultar *experts* para lhe fornecerem as informações necessárias (*right to rely*), mas afirma que é do administrador a

decisões tomadas com base em sistemas de inteligência artificial. In: FRAZÃO, Ana; MULHOLLAND, Caitlin (Coord.). **Inteligência artificial e Direito: ética, regulação e responsabilidade**. 2. ed. rev. atual. e ampl. São Paulo: Thomson Reuters Brasil, 2020. p. 86. *E-book*.

⁹ FRAZÃO, Ana. Responsabilidade civil dos administradores de sociedades empresárias por decisões tomadas com base em sistemas de inteligência artificial. In: FRAZÃO, Ana; MULHOLLAND, Caitlin (Coord.). **Inteligência artificial e Direito: ética, regulação e responsabilidade**. 2. ed. rev. atual. e ampl., p. 481-525, São Paulo: Thomson Reuters Brasil, 2019. p. 494-495.

responsabilidade por escolher o sistema e por avaliar sua confiabilidade, o que deve ser feito com prudência, cautela e máxima diligência. Mesmo com todas as dificuldades técnicas para compreender a lógica algorítmica, tais como a opacidade dos sistemas, é papel do administrador avaliar sua acurácia, robustez, bem como acompanhar seus riscos e resultados.

Este entendimento vai ao encontro do que foi estudado no capítulo anterior, com relação ao dever de se informar e ao dever de se fiar no assessoramento de terceiros. Como lá visto, mesmo que o administrador busque se amparar em opiniões de especialistas, ele não pode simplesmente replicá-las sem assegurar uma série de requisitos e cautelas antes e depois do recebimento das orientações.

A doutrina especializada e os precedentes da CVM exigem que a consulta à especialista seja reconhecida como diligência quando o assessor consultado (i) tiver reconhecida experiência no tema avaliado pela administração, (ii) seja subsidiado com todos os dados e informações pertinentes para alicerçar suas conclusões, (iii) tenha suas orientações validadas pelo administrador, (iv) esgote todos os possíveis sinais de alerta, (v) seja idônea e de boa-fé e, quando, após tudo isso, (vi) for seguida pelos administradores.

Reforça-se que, ainda que toda cautela necessária tenha sido adotada na eleição da IA a ser utilizada, o seu monitoramento e fiscalização merecem a mesma atenção, sob pena de se descumprir subdeveres do dever de diligência:

Disso decorre uma primeira conclusão: a de que a transferência total de processos decisórios para os sistemas de inteligência artificial que podem causar danos à sociedade, aos sócios ou a terceiros, sem nenhum controle, acompanhamento ou intervenção humana, não é compatível com o dever de diligência, diante dos conhecidos riscos que podem daí decorrer. A delegação é ainda mais grave se levar à transferência completa, ainda que informalmente, da própria competência de gestão para uma máquina. Assim, a utilização da inteligência artificial pelos administradores e órgãos administrativos deve ocorrer com os cuidados e responsabilidades devidos e o monitoramento

constante, o qual dependerá de uma série de circunstâncias e de características da sociedade empresária, inclusive o seu porte, tipo de atividade e grau de risco.¹⁰

Por conseguinte, Frazão adverte que o administrador deve criar organização que possibilite a testagem periódica dos resultados da utilização da IA, para que se identifiquem resultados indesejados, especialmente aqueles envolvendo ilícitos e/ou discriminações negativas. Conclui-se, portanto, que o dever de diligência dos administradores pode ser traduzido essencialmente na obrigação de agir informado e de criar organização compatível com o risco da inteligência artificial que deseja utilizar, o que será avaliado neste tópico.

Deste modo, entende-se que os deveres e subdeveres inerentes ao dever de diligência devem ser empregados pelos administradores de empresas que desejam utilizar aplicações tecnológicas com IA, tornando-se, portanto, necessário estabelecer e compreender como construir um *iter* decisório que seja posteriormente reconhecido como diligente, mesmo em caso de falhas da ferramenta ou de produção de resultados indesejados, como pode ser a discriminação algorítmica.

Portanto, a seguir, a pesquisa dedicar-se-á a examinar o *compliance*, tendo em vista que este termo foi mencionado no capítulo anterior como um instrumento mínimo a ser adotado para dar concretude ao dever de diligência, fazendo com que sua implementação se torne um verdadeiro dever para o administrador.

Ao final, buscar-se-á compreender como tais estruturas e ferramentas, criadas inicialmente para prevenir conflitos e danos de outras naturezas, podem ser úteis para garantir um uso de IA mais cauteloso e diligente, a ponto de mitigar ou até evitar a ocorrência de discriminação algorítmica.

4.1 Compliance

Nas palavras do Ministro do Supremo Tribunal Federal Luis Roberto Barroso, o termo

¹⁰ FRAZÃO, Ana. Responsabilidade civil dos administradores de sociedades empresárias por decisões tomadas com base em sistemas de inteligência artificial. In: FRAZÃO, Ana; MULHOLLAND, Caitlin (Coord.). **Inteligência artificial e Direito: ética, regulação e responsabilidade**. 2. ed. rev. atual. e ampl., p. 481-525, São Paulo: Thomson Reuters Brasil, 2019. p. 509.

Compliance significa a observância das normas legais e regulamentares aplicáveis, bem como dos valores éticos gerais, dos códigos de conduta específicos de determinado ramo da atividade e das expectativas legítimas da sociedade. O termo também abarca os programas de incentivos a tais condutas.¹¹

O Ministro cita as diversas leis brasileiras que preveem incentivos para a adoção de estruturas de *compliance*¹², tais como a Lei 12.846/2013 (Lei Anticorrupção), regulada pelo Decreto 8.420/2015, bem como a Lei 13.303/2016 (Lei das Estatais). Mais recentemente, ainda pode-se acrescentar a Lei Geral da Proteção de Dados, sem esquecer da presença deste instituto no projeto de Lei n.º 21/2020 e aglutinados, que discutem a criação de um Marco Legal para Inteligência Artificial no Brasil (notadamente, no Capítulo IV - DA GOVERNANÇA DOS SISTEMAS DE INTELIGÊNCIA ARTIFICIAL, constante na versão produzida pela Comissão de Juristas e que está em análise no Senado Federal).

Apesar de o termo *compliance* ter se popularizado no Brasil como forma de evitar a corrupção, hoje já se reconhece sua pertinência para mitigação de riscos em diversas áreas dentro de uma empresa, tais como criminal, antifraude, proteção de dados, trabalhista, consumidor, ambiental, entre outros, a depender da atividade empresarial, porte e recursos da empresa.

Segundo Milena Donato Oliva e Rodrigo da Guia Silva, o *compliance* é uma forma de autovigilância para as pessoas jurídicas, por meio de mecanismos de autorregulação e autorresponsabilidade. Neste sentido, alertam que não há um modelo ideal de *compliance*, sendo possível adequá-lo às particularidades de cada empresa, de modo que ele seja efetivo e atinja os objetivos de evitar infrações, ou, quando não conseguir, detectá-las e contribuir para a sua correção.¹³

¹¹ FRAZÃO, Ana; CUEVA, Ricardo Villas Bôas (Org.). **Compliance: perspectivas e desafios dos programas de conformidade**. 1. ed. Belo Horizonte: Forum, 2018. v. 1. 527p.

¹² Registra-se que tais leis não foram objeto de exame deste estudo e são mencionadas apenas como reforço retórico para exemplificar as diversas legislações que incentivam a adoção de instrumentos que visem a conformidade com as leis e regulamentos, evidenciando a relevância do *compliance* no ambiente empresarial.

¹³ OLIVA, Milena Donato; SILVA, Rodrigo da Guia. Origem e evolução histórica do compliance no Direito brasileiro. In: FRAZÃO, Ana; CUEVA, Ricardo Villas Bôas (Org.). **Compliance: perspectivas e desafios dos programas de conformidade**. 1. ed. Belo Horizonte: Forum, 2018. v. 1. 527p. p. 29-51.

Os autores fazem um levantamento do arcabouço legal e regulatório que menciona o *compliance*, com destaque para as exigências feitas pela CVM e pelo Novo Mercado da B3, o que demonstra a relevância destes programas para a credibilidade e reputação das empresas perante clientes e investidores. Também evidenciam que o *compliance* é importante na prevenção e reparação de infrações à ordem jurídica, traçando relações entre seus objetivos e os objetivos da responsabilização civil (como o papel preventivo), apesar das profundas distinções entre os dois institutos. Nesta medida, indicam que:

Os mecanismos de compliance, por sua vez, destinam imediatamente a uma autêntica prevenção de danos, atuando no sentido de mitigarem os riscos de ilícitos associados as atividades do agente econômico, sem prejuízo da importância do programa de compliance na reparação de danos e no retorno ao cenário de legalidade se não se tiver logrado evitar a infração.

[...]

O compliance, ao revés, é modulado pelo agente econômico a partir das especificidades do seu negócio, estabelecendo sanções disciplinares dentro dos limites legalmente admitidos, bem como cooperando com o Poder Público sempre que necessário. [...] é mais efetivamente vocacionado à prevenção de infrações e de danos, tendo por função primordial evitar a ocorrência de violações às normas legais aplicáveis, bem como ao Código de Ética próprio de cada pessoa jurídica.¹⁴

Vale destacar que um *compliance* efetivo pode significar o afastamento de responsabilizações de natureza administrativa ou, ao menos, uma atenuante de sanções, eis que atesta a "seriedade do agente econômico". Com isso, também possibilita mais negócios e maior inserção no mercado. Portanto, os incentivos para a sua implementação não decorrem só de leis, mas "dos parceiros comerciais, consumidores, funcionários, porque o *compliance* busca assegurar ambiente corporativo sério, saudável e comprometido com a legalidade."¹⁵

¹⁴ OLIVA, Milena Donato; SILVA, Rodrigo da Guia. Origem e evolução histórica do compliance no Direito brasileiro. In: FRAZÃO, Ana; CUEVA, Ricardo Villas Bôas (Org.). **Compliance: perspectivas e desafios dos programas de conformidade**. 1. ed. Belo Horizonte: Forum, 2018. v. 1. 527p. p. 41.

¹⁵ *Ibidem*, p. 48.

Já Ricardo Villas Bôas Cueva define o *compliance* como:

instrumentos de governança corporativa tendentes a garantir que as políticas públicas sejam implantadas com maior eficiência. Compõem-se de rotinas e práticas concebidas para prevenir riscos de responsabilidade empresarial decorrentes do descumprimento de obrigações legais ou regulatórias. [...] voltam-se para a mudança de comportamento, por meio de padrões de conduta a serem observados e monitorados pelas empresas [...] a fim de evitar o cometimento de ilícito.

De modo mais geral, pode-se entender o *compliance* não apenas como observância de comandos legais e regulatórios, mas também como o cumprimento de outras exigências, tais como normas éticas, padrões de conduta fixados no seio das organizações e expectativas dos stakeholders.¹⁶

O autor leciona que os programas são aliados estatais na regulação da atividade empresarial, tendo em vista a dificuldade que se tem de acompanhar seu desenvolvimento e conformidade no atual contexto de complexidade social e avanço tecnológico, o que a doutrina denomina como autorregulação regulada ou correção. Inclusive, Cueva pontua as críticas a esse entendimento, que pode desvirtuar o foco do *compliance* – que é evitar ilícitos.

Na visão de Cueva,

As funções dos programas de *compliance* compreendem a proteção da empresa, dos administradores e empregados, a orientação e a informação dos empregados, a garantia de qualidade, o fomento à inovação e o monitoramento. A função protetiva corresponde ao objetivo primordial dos programas de conformidade, ou seja, evitar infrações de regras por meio de medidas organizacionais preventivas. Procura-se proteger a empresa, os administradores e os empregados de sanções civis, administrativas e penais. Com isso, a

¹⁶ CUEVA, Ricardo Villas Bôas. Funções e finalidades dos programas de *compliance*. In: CUEVA, Ricardo Villas Bôas; FRAZÃO, Ana (Coord.). **Compliance**: perspectivas e desafios dos programas de conformidade. Belo Horizonte: Fórum, 2018. p. 53-69.

empresa tende a evitar não apenas ônus financeiros ou limitações às suas atividades, mas também danos à sua reputação.¹⁷

Além do descrito acima, o autor refere sobre as funções de: orientar e informar a administração e os empregados; treinamento e informação; garantir a qualidade e o fomento a inovação; a adequação e melhorias a partir dos *feedbacks* recebidos de clientes; o monitoramento e supervisão contínuos. Entende que para além de um instrumento de governança corporativa, os programas protegem a empresa, seus administradores e seus empregados, na medida em que impactam significativamente na avaliação de responsabilidade da empresa e de seus administradores no caso de ocorrência de alguma infração, podendo ensejar atenuação ou redução de sanções.

É em virtude destas funções de evitar danos significativos para as organizações empresariais que Lucas defende a adoção de estruturas de *compliance* não como uma opção, mas uma decorrência obrigatória para se efetivar o cumprimento do dever de diligência dos administradores.¹⁸

Neste sentido, Frazão também sustenta que é dever do administrador adotar normas de *compliance* (prevenção de riscos, canais de comunicação, fluxos de testagem e auditoria) durante o ciclo da vida da inteligência artificial. A autora ainda elenca princípios a serem observados na criação destes fluxos, com base na LGPD e no Guia da Comissão Europeia (abril/19). Destaca que as responsabilidades em caso de danos ocasionados a terceiros por decisões tomadas com base em inteligência artificial utilizada ou desenvolvida por sociedade anônimas:

- a) Responde a empresa, na forma da lei cabível (CLT, CDC, CC, LGPD, etc.);
- b) Responde solidariamente a pessoa do administrador, quando agir com dolo ou culpa, ou quando violar lei ou estatuto.

Ainda, conforme a autora, o critério sugerido para avaliação da existência de culpa é a observância (ou não) do cumprimento do dever de diligência, ou seja, avaliar

¹⁷ CUEVA, Ricardo Villas Bôas. Funções e finalidades dos programas de compliance. In: CUEVA, Ricardo Villas Bôas; FRAZÃO, Ana (Coord.). **Compliance: perspectivas e desafios dos programas de conformidade**. Belo Horizonte: Fórum, 2018. p. 53-69.

¹⁸ LUCAS, Laís Machado. **Programas de integridade nas sociedades anônimas: implementação como conteúdo do dever de diligência dos administradores**. Porto Alegre: Livraria do Advogados, 2021.

se o administrador teria condições de agir de maneira mais zelosa. Ela também destaca que, como já visto em capítulo anterior, o dever de diligência é uma obrigação de meio, não de resultado. O foco na atribuição da responsabilidade não deve ser a dimensão do dano causado, mas avaliar se todas as cautelas possíveis foram tomadas para evitá-lo ou mitigá-lo, o que atrai ainda mais importância para a existência de um *compliance* sério.

Assim, além de configurar um instrumento necessário para efetivação do dever de diligência, o *compliance* também é o que viabiliza a conformação da empresa ao ordenamento jurídico em que está inserida, prezando pelo cumprimento dos deveres legais.

Gustavo Machado Gonzales e Bruno Tostes Corrêa inclusive defendem que há um dever implícito de legalidade imposto aos administradores.¹⁹ Mencionam que, apesar de não ser explícito na legislação, tal dever é amplamente reconhecido em diversos países e aparece em decisões da CVM, mesmo que não haja menção expressa. Os autores elencam cinco fundamentos para a sua existência:

- a) O postulado da unidade do ordenamento jurídico: sendo sujeitos de direito, as sociedades e seus administradores devem estar em conformidade com a ordem jurídica;
- b) Os sujeitos de direito têm interesse que as normas jurídicas sejam cumpridas por todos os atores sociais (de modo que se possam estabelecer relações). Assim, para que seja efetivado o cumprimento das normas, é necessário reconhecer o dever/responsabilidade dos administradores das sociedades de perseguir este objetivo;
- c) Necessidade de tutelar os interesses jurídicos que incidem sobre as sociedades, como os penais, ambientais, concorrenciais etc.: os administradores devem assegurar que as sociedades estejam de acordo com as leis;
- d) A exploração da atividade econômica só é legítima quando estiver de acordo com as regras do jogo. A obtenção de lucro pressupõe a observância de normas jurídicas;

¹⁹ GONZALES, Gustavo Machado; CORRÊA, Bruno Tostes. Dever de legalidade dos administradores de sociedades. **Revista Semestral de Direito Empresarial**, n. 30 (janeiro/junho 2022), Rio de Janeiro: Renovar, 2007.

e) Considerando que a desobediência às leis pode ocasionar prejuízos à sociedade (ex. sanções diversas como multas, indenizações, danos à reputação), causando a diminuição do patrimônio social, os administradores devem assegurar que isto não aconteça.

Após tais considerações, os autores demonstram como o dever de legalidade dos administradores nasceu em decisões norte-americanas, sendo, após, reconhecido explicitamente pelo direito alemão e outros países europeus, como Espanha e Portugal. Em seus estudos, mencionam como as decisões estrangeiras demonstram ser dever dos administradores a criação de estruturas/mecanismos/controles internos adequados de monitoramento das atividades da companhia, evitando descumprimentos legais.

Assim, definem que o dever de legalidade é:

o dever de legalidade também impõe que o administrador promova ou assegure o cumprimento, pela sociedade, de todas as normas que recaem sobre ela como sujeito de direito autônomo em suas relações com terceiros.

[...]

o dever de legalidade, que obriga o administrador, para além de cumprir todos os deveres incidentes diretamente sobre ele, a “diligenciar ou assegurar que a sociedade, ela própria, se comport[e] correctamente face à ordem jurídica, cumprindo aquilo que a ordem jurídica dela exige.”²⁰

Os autores refletem que em decorrência deste dever de legalidade há:

um dever de criar mecanismos de controle que garantam o cumprimento da lei por parte da sociedade, geralmente conhecidos pelo termo compliance. [...]

Este conteúdo do dever de legalidade, referente à necessidade de constituição de mecanismos de controles internos, pode ser visto, inclusive, no princípio 5 do Código de Governança Corporativa alemão (Deutscher Corporate Governance Kodex – DCGK), o qual prescreve que o Vorstand da companhia aberta alemã

²⁰ GONZALES, Gustavo Machado; CORRÊA, Bruno Tostes. Dever de legalidade dos administradores de sociedades. **Revista Semestral de Direito Empresarial**, n. 30 (janeiro/junho 2022), Rio de Janeiro: Renovar, 2007. p. 30.

deve garantir que todas as normas e políticas internas sejam cumpridas e que se esforçará para obter seu cumprimento por parte da empresa.²¹

Analisando sob o ordenamento jurídico brasileiro, os autores compreendem que a interpretação mais coerente com o sistema societário impõe o reconhecimento de um dever de legalidade, afirmando que “[...] o dever de legalidade dos administradores sociais é uma consequência natural da prevalência da ordem jurídica e da necessidade de cumprimento das normas pelas sociedades, ainda que inexistam textos que contenha explicitamente [...]”²² Extraem de entendimentos doutrinários e julgados da CVM trechos que evidenciam que há um acolhimento implícito sobre a necessidade de os administradores respeitarem a legislação e reforçam o dever de implementar programas de *compliance*:

hoje, o dever de monitoramento, como consectário do dever de diligência, requer também falar-se em dever de constituição de estruturas adequadas de gestão e controle, a fim de monitorar de forma adequada os negócios da companhia. Ou, ainda, a proposição de que a implementação de programas de integridade estaria dentro da órbita do dever de diligência.²³

Assim, os autores concluem pela existência de duas dimensões do dever de legalidade:

- (a) a primeira se refere ao dever do administrador de não tomar parte, como autor ou mero partícipe, de um ato que leve a sociedade a violar o ordenamento jurídico;
- (b) a segunda, por outro lado, diz respeito ao dever do administrador de diligenciar para que a sociedade não viole as normas jurídicas – é nesta dimensão que surge o dever de constituir sistemas de controles internos (dever de controle da legalidade).²⁴

²¹ *Ibidem*, p. 31.

²² *Ibidem*, p. 44.

²³ GONZALES, Gustavo Machado; CORRÊA, Bruno Tostes. Dever de legalidade dos administradores de sociedades. **Revista Semestral de Direito Empresarial**, n. 30 (janeiro/junho 2022), Rio de Janeiro: Renovar, 2007. p. 50.

²⁴ *Ibidem*, p. 59.

Consoante exposto neste tópico, verifica-se que a implementação de programas de *compliance* não é mais vista como uma liberalidade das companhias, mas sim como um dever, o que se interpreta tanto a partir da análise do dever de diligência, quanto da análise de um dever de cumprir as normas legais – que, no caso desta pesquisa, é a de tratar todos com igualdade, sem discriminação, ainda que tais ações discriminatórias sejam realizadas por algoritmos sem qualquer intencionalidade.

Assim, no item seguinte, a pesquisa se dedicará a investigar se a adoção de programas de *compliance* no ciclo de vida de inteligência artificial atinge o objetivo de reduzir a discriminação algorítmica e configura a diligência esperada de um administrador para evitar ou mitigar danos ocasionados a terceiros e, por reflexo, à reputação da própria empresa.

4.2 Entendendo as fontes de vieses e suas possíveis soluções: o *compliance* pode evitar ou mitigar a discriminação algorítmica?

Antes de avançar a pesquisa abordando as ferramentas hábeis a mitigar a discriminação praticada por IA, é preciso entender em qual parte da mecânica da formação dos algoritmos que ela ocorre, para, deste modo, poder impor as medidas de correção – e isso se elas realmente existirem, pois, do contrário, só caberá educar a população usuária a ter maior cuidado ou então proibir de vez o uso de IA em determinadas situações.

Grande parte das pesquisas jurídicas que abordam inteligência artificial focam na definição de princípios éticos e legais que precisam ser respeitados para se ter uma IA benéfica.²⁵

Outro argumento muito enfrentando é a respeito da opacidade dos algoritmos, também chamados de "*black boxes*", tendo em vista que suas regras de programação, além de chegarem a conclusões por meio de metodologias que nem

²⁵ CORRÊA, Nicholas Kluge et al. Worldwide AI ethics: a review of 200 guidelines and recommendations for AI governance. *Veritas*, Porto Alegre, 2022. Disponível em: <http://dx.doi.org/10.48550/arXiv.2206.11922>. Acesso em: 14 jan. 2023.

sempre são compreensíveis, normalmente estão protegidas por segredo comercial, o que torna "opaca" a visão acerca de seu funcionamento.

Então, apesar de muitas das pesquisas apontarem reais necessidades a serem discutidas para que a IA não cause danos a alguém, não é frequente encontrar alguma indicação de como contornar tais problemas e aplicar princípios e valores éticos de forma efetiva.

Todavia, há estudos na área da computação que relatam melhores práticas a serem utilizadas pelas empresas no desenvolvimento de uma IA durante seu processo de "aprendizagem", as quais, quando implementadas, podem evitar ou reduzir seus efeitos colaterais discriminatórios – o que demonstra que um programa de *compliance* que impusesse o cumprimento de certos requisitos nestas etapas poderia ser efetivo para evitar ou mitigar a discriminação algorítmica.

Um exemplo destes estudos é a pesquisa feita por Harini Suresh e John Guttag no Massachusetts Institute of Technology (MIT), que identificou que os vieses algorítmicos podem ocorrer em diferentes fases do ciclo de vida da aprendizagem de máquina. Também foi possível constatar que a solução dada para um tipo de viés pode ser danosa para solucionar outro tipo e que, nem sempre, os problemas enfrentados têm origem nos dados utilizados para treinamento da IA:

Machine learning (ML) is increasingly used to make decisions that affect people's lives. Typically, ML algorithms operate by learning models from existing data and generalizing them to unseen data. As a result, problems can arise during the data collection, model development, and deployment processes that can lead to distinct types of harmful downstream consequences. [...]

A common refrain is that undesirable behaviors of ML systems happen when "data is biased." Indeed, a recent comment by a prominent ML researcher to this end set off a heated debate—not necessarily because the statement "data is biased" is false, but because it treats data as a static artifact divorced from the process that produced it. This process is long and complex, grounded in historical context and driven by human choices and norms. Understanding the implications of each stage in the data-generation process can reveal more direct and meaningful ways to prevent or address harmful downstream consequences that can be masked by overly broad terms like "biased data."

*Moreover, it is important to acknowledge that not all problems should be blamed on the data. The ML pipeline involves a series of choices and practices, from model definition to user interfaces used upon deployment. Each stage involves decisions that can lead to undesirable effects. For an ML practitioner working on a new system, it is not straightforward to identify if and how problems might arise. Even once identified, it is not clear what the appropriate application- and data-specific mitigations might be, or how they might generalize over factors such as time and geography.*²⁶

Em suma, os cientistas partem da premissa de que é mais produtivo identificar as causas dos potenciais danos que podem surgir em um processo de aprendizagem de máquina, incentivando soluções apropriadas para cada aplicação, do que tentar traduzir em noções amplas e principiológicas o que é justo.

Assim, os autores examinaram casos famosos da literatura sobre discriminação algorítmica, muitos dos quais citados no primeiro capítulo, para diagnosticar em que fase do ciclo de vida de aprendizagem de máquina a distorção

²⁶ Em tradução livre: O aprendizado de máquina (ML) é cada vez mais usado para tomar decisões que afetam a vida das pessoas. Normalmente, os algoritmos de ML operam aprendendo modelos a partir de dados existentes e generalizando-os para dados não vistos. Como resultado, podem surgir problemas durante a coleta de dados, desenvolvimento de modelos e processos de implantação que podem levar a diferentes tipos de consequências prejudiciais para o sistema. [...].

Um bordão comum é que comportamentos indesejáveis dos sistemas ML acontecem quando "os dados são tendenciosos". De fato, um comentário recente de um pesquisador proeminente do ML para este fim desencadeou um debate acalorado - não necessariamente porque a afirmação "dados são tendenciosos" é falsa, mas porque trata os dados como um artefato estático dissociado do processo que os produziu. Este processo é longo e complexo, fundamentado no contexto histórico e impulsionado por escolhas e normas humanas. A compreensão das implicações de cada etapa do processo de geração de dados pode revelar formas mais diretas e significativas de prevenir ou lidar com as consequências prejudiciais ao sistema que podem ser mascaradas por termos excessivamente amplos como "dados tendenciosos".

Além disso, é importante reconhecer que nem todos os problemas devem ser imputados aos dados. O ciclo de aprendizagem de máquina envolve uma série de escolhas e práticas, desde a definição do modelo até as interfaces de usuário utilizadas na implantação. Cada etapa envolve decisões que podem levar a efeitos indesejáveis. Para um profissional do ML que trabalha em um novo sistema, não é simples identificar se e como os problemas podem surgir. Mesmo uma vez identificados, não está claro quais podem ser as mitigações apropriadas para aplicações e dados específicos, ou como elas podem se generalizar sobre fatores como tempo e geografia. SURESH, Harini; GUTTAG, John. Understanding potential sources of harm throughout the machine learning life cycle. **MIT Case Studies in Social and Ethical Responsibilities of Computing**, 10 ago. 2021. E-book. Disponível em: <https://mit-serc.pubpub.org/pub/potential-sources-of-harm-throughout-the-machine-learning-life-cycle/release/2?from=7166&to=7866>. Acesso em: 13 nov. 2022.

ocorrida pode ter sido gerada, identificando pelo menos sete fontes de vieses, resumidas por Nicholas Corrêa, conforme abaixo:

- **Vieses históricos:** este tipo de problema ocorre pelo fato de que o nosso próprio mundo, como é ou foi, é falho. Assim, mesmo que o modelo seja uma representação perfeita do ambiente, ele ainda poderá gerar danos, pois representa um ambiente imperfeito. Por exemplo, Brown et al. (2020, p. 36-37) reportam que seu modelo de linguagem (GPT-3) associa adjetivos pejorativos, sexistas e misóginos com mais frequência há [sic] mulheres do que homens (i.e., uma reflexão dos textos, e cultura, que encontramos pela internet);
- **Vieses de representação:** este problema ocorre quando os dados utilizados para o treinamento do modelo não representam a população ou ambiente em que o modelo irá agir. Quando treinando um modelo para prever o desenvolvimento de diabetes tipo 2, talvez não haja dados o suficiente para representar todos os possíveis grupos de interesse (homens, mulheres, idosos, crianças, etc.). Ou, um software de reconhecimento de imagem utilizado por um carro autônomo treinado em ambientes urbanos, pode operar de forma falha quando operando em regiões rurais;
- **Vieses de medição:** quando escolhemos características (features) para serem utilizadas por algum resultado (e.g., IMC, peso, altura, histórico familiar), estamos supondo que tal característica/quantidade é representativa daquilo que queremos prever ou classificar (diabetes tipo 2). Mas esse nem sempre é o caso, pois tais "proxies" podem ser apenas aproximações de uma realidade mais complexa. Por exemplo, se nosso modelo atribui muito peso a variável "IMC" para a tarefa de predição de diabetes tipo 2, sujeitos com um grande volume muscular (e.g., fisiculturistas) poderão ser classificados falsamente como potenciais desenvolvedores de diabetes tipo 2. "QI" (i.e., quociente de inteligência) pode não ser um bom parâmetro para se avaliar sucesso acadêmico, que muitas vezes depende de outros fatores difíceis de serem mensurados (e.g., motivação, capacidade de se relacionar, habilidades organizacionais);
- **Vieses de agregação:** este tipo de problema ocorre quando grupos diferentes são unidos em um único conjunto de dados. Contudo, o modelo treinado não atua eficientemente com algum (ou nenhum) dos grupos. Por exemplo, sabe-se que homens têm duas vezes mais chances de ter um ataque cardíaco ao longo da

vida. Um modelo treinado com um conjunto de dados misto (i.e., sem diferenciar o sexo das amostras) para prever a chance de um paciente ter um ataque cardíaco, pode vir a ser ineficiente para com algum dos sexos (ou ambos). Idealmente, modelos devem ser treinados para atender grupos específicos (quando necessário);

- Vieses de aprendizagem: a escolha da função de perda (e.g., erro médio quadrático, entropia cruzada binária, entropia cruzada categórica) e métrica de performance (e.g., acurácia, precisão, recall, AUC) pode influenciar o tipo de saída que nosso modelo gera, e como interpretamos sua performance. Por exemplo, se temos uma aplicação para qual classificações negativas falsas podem gerar um “grande custo” (e.g., falso negativo para HIV), talvez não devamos utilizar acurácia para medir sua performance, mas sim recall;
- Vieses de avaliação: nem sempre os dados utilizados na fase de testagem (ou teste-padrão/benchmark) representam uma boa métrica de avaliação para o domínio em que o modelo será implantado. Por exemplo, você pode ter desenvolvido um modelo de reconhecimento facial com uma excelente performance em sua fase de teste. Contudo, o benchmark que você utilizou tem uma baixa representação da população parda (e.g., 4%), e o modelo irá atuar em um domínio onde grande parte da população é parda (e.g., Brasil);
- Vieses de implantação: este tipo de problema ocorre quando o modelo é utilizado de forma diferente, ou além, daquilo que foi originalmente desenvolvido para fazer. Muitos dos modelos criados por aprendizagem de máquina não são “totalmente autônomos”, mas se encontram como parte de um processo sociotécnico onde intenções e desejos humanos fazem parte. Por exemplo, sistemas de avaliação de risco são utilizados no sistema penal americano para prever a probabilidade de uma pessoa cometer um crime futuro (i.e., reincidência criminal). Contudo, uma instanciação perversa desta ferramenta seria a utilizar para determinar a duração de uma sentença com base no risco provável de reincidência.²⁷

²⁷ COLLINS (2018) *apud* CORRÊA, Nicholas Kluge. **Ética e segurança da inteligência artificial: ferramentas práticas para se criar "bons" modelos**. Dez. 2021. Disponível em: https://www.researchgate.net/publication/357032340_Etica_e_Seguranca_da_Inteligencia_Artificial_ferramentas_praticas_para_se_criar_bons_modelos. Acesso em: 13 nov. 2022.

Deste modo, seguindo a ordem demonstrada acima, denota-se que os *vieses históricos* normalmente surgem na fase de coleta de dados, o que ocorre mesmo quando os dados estejam perfeitamente medidos e englobem uma amostra ideal. Contudo, como reflete o mundo (e seus defeitos) leva a um modelo que produz resultados discriminatórios. Esta situação pôde ser vista no exemplo descrito anteriormente, que envolveu a seleção de CEOs ou a reprodução de conceitos machistas em pesquisas no Google.

Segundo os autores, este último, em específico, ocorre, pois aplicações que envolvem processamento de linguagem natural atribuem significação a palavras de acordo com textos disponíveis na *web*, tais como notícias, Wikipedia etc. Por isso, tal técnica pode refletir preconceitos humanos históricos, tendo em vista que não distingue o contexto em que a palavra está inserida, podendo replicá-la mesmo que seja negativa.

Outro caso envolvendo problema semelhante foi o da Tay, a IA desenvolvida pela Microsoft para interagir com jovens de 18 a 24 anos no Twitter e que em menos de vinte quatro horas publicou conteúdos racistas, machistas e defendeu o nazismo. Na ocasião, Peter Lee, vice-presidente da empresa, manifestou-se a respeito do insucesso de Tay:

*AI systems feed of both positive and negative interactions with people. In that sense, the challenges are just as much social as they are technical. We will do everything possible to limit technical exploits but also know we cannot fully predict all possible human interactive misuses without learning from mistakes. To do AI right, one needs to iterate with many people and often in public forums.*²⁸

²⁸ Tradução livre: Os sistemas de IA se alimentam de interações positivas e negativas com as pessoas. Nesse sentido, os desafios são tanto sociais quanto técnicos. Faremos todo o possível para limitar as explorações técnicas, mas também sabemos que não podemos prever totalmente todos os possíveis usos incorretos interativos humanos sem aprender com os erros. Para fazer a IA corretamente, é preciso interagir com muitas pessoas e frequentemente em fóruns públicos. Devemos entrar em cada um com muita cautela e, finalmente, aprender e melhorar, passo a passo, e fazer isso sem ofender as pessoas no processo. Permaneceremos firmes em nossos esforços para aprender com esta e outras experiências enquanto trabalhamos para contribuir para uma Internet que represente o melhor, não o pior, da humanidade. LEE, Peter. Learning from Tay's introduction. **Microsoft**, 25 mar. 2016. Disponível em: <https://blogs.microsoft.com/blog/2016/03/25/learning-tays-introduction> sm.00000gjdppwwcfcus11t6oo6dw79gw. Acesso em: 18 nov. 2022.

Conforme Suresh e Guttag, reconhecer um viés histórico requer a compreensão de como a opressão se manifestou ao longo do tempo para que, a partir daí, se possam definir as estratégias de mitigação.²⁹

Com relação ao viés de *representação*, que pode ocorrer durante a preparação dos dados coletados, este poderá se manifestar quando a amostra de desenvolvimento representa apenas uma parcela da população, e, conseqüentemente, não consegue generalizar sua aprendizagem.

Este exemplo foi visto no caso denunciado por Joy Buolamwini, que identificou falhas nas aplicações de reconhecimento facial utilizada por grandes empresas quando aplicadas em mulheres de pele escura, pois os bancos de dados de imagens públicas, utilizados para treinamento dos algoritmos, continham poucas imagens desta parcela da população, o que impediu a IA de “aprender” a reconhecer rostos com aquelas características.³⁰ Como no caso de Buolamwini, corrigir as amostras contidas nos bancos de dados, com o intuito de torná-los mais representativos, pode ser uma forma de contornar este viés.

Já o viés de *medição* pode ocorrer na utilização de *proxies* para medir determinada situação complexa, pois, segundo os autores, “*The proxy is an oversimplification of a more complex construction.*”³¹

O problema da medição também pode ocorrer quando o método varia entre grupos, como mencionado nos casos envolvendo COMPAS e policiamento preditivo. Um dos critérios elencados para julgar o grau de periculosidade e reincidência de um réu era seu bairro, seus vizinhos, contatos com outros crimes, drogas ou álcool. Contudo, viu-se que o policiamento mais ostensivo com negros e latinos gerava mais fichas criminais desta parcela da população do que de brancos de bairro nobre que não eram alvo de tantas fiscalizações policiais – reforçando, equivocadamente, a ideia de que bairros predominantemente habitados por negros mereciam mais

²⁹ SURESH, Harini; GUTTAG, John. Understanding potential sources of harm throughout the machine learning life cycle. **MIT Case Studies in Social and Ethical Responsibilities of Computing**, 10 ago. 2021. Disponível em: <https://mit-serc.pubpub.org/pub/potential-sources-of-harm-throughout-the-machine-learning-life-cycle/release/2?from=7166&to=7866>. Acesso em: 13 nov. 2022. *E-book*.

³⁰ BUOLAMWINI, Joy, GEBRU, Tinmit. “Gender shades: intersectional accuracy disparities in commercial gender classification.” **Proceedings of Machine Learning Research**, v. 81, n. 1, p. 1–15, 2018. Conference on Fairness, Accountability, and Transparency.

³¹ Em tradução livre: O *proxy* é uma simplificação exagerada de uma construção mais complexa. SURESH; GUTTAG, *loc. cit.*

policiamento e maiores condenações, criando um ciclo de *feedback* ou efeito *looping* em razão da medição desigual. Neste caso, a revisão dos *proxies* eleitos deve ser uma medida implementada para redução deste viés, focando especialmente em variáveis que contenham nexos de causalidade com o resultado perseguido pela IA, e não de correlação.

O quarto viés mencionado, o viés de *agregação*, ocorre na etapa de desenvolvimento do modelo. Pode ocorrer quando se desconsideram características específicas em prol de um enquadramento mais genérico, buscando a simplificação ou padronização do modelo. Este viés foi diagnosticado, por exemplo, na análise de aplicações de processamento de linguagem, em que palavras ou frases que poderiam representar agressões, em determinados locais significam apenas letras de um *rapper*, pois elas foram enquadradas em grupo geral (ex. "xingamentos"), sem análise de contexto. Estes casos, portanto, exigem constante aprimoramento do modelo, para que possa conjugar novas variáveis em sua análise, reduzindo-se as generalizações ao considerar mais dados do contexto.

O viés de *aprendizagem* pode ocorrer na avaliação do desempenho do modelo, ao se escolher a métrica que irá avaliar a performance da aplicação, o que também pode depender do seu objetivo. Por exemplo, se for um modelo que avalia a capacidade de adimplência de quem solicita um empréstimo, os candidatos se preocuparão com os falsos negativos (ou seja, chance de ter um empréstimo negado quando poderia ser concedido) enquanto os bancos/credores irão se preocupar em avaliar os falsos positivos (quando o sistema indica a aprovação de empréstimos a devedores com alto risco de inadimplência). Assim, a depender do interessado no desenvolvimento do modelo, pode-se atribuir mais importância para uma circunstância ou outra. Como se vê, a correção deste viés envolve a revisão dos critérios eleitos para avaliação do modelo.

O viés de *avaliação* pode ocorrer quando se elege uma métrica para avaliação de desempenho de um algoritmo sem saber seu real desempenho – por exemplo, no caso de Joy Buolamwini, diagnosticou-se que a IA de reconhecimento facial apresentava uma alta taxa de erro quando provocadas a reconhecer rostos de mulheres de pele escura. Todavia, isto, provavelmente ocorreu pois o modelo continha poucas amostras desta população para treinar e não porque era um mau

modelo, necessariamente. Talvez, se tivesse recebido mais amostras de treinamento, poderia apresentar um desempenho excelente.

O viés de *implantação* ocorre quando o modelo, desenvolvido para determinada função, é utilizado para outra. O exemplo a ser considerado também pode ser o do caso COMPAS, que foi desenvolvido para fins de indicar o grau de reincidência na avaliação da fiança e acabou sendo utilizado para embasar a dosimetria da pena dos condenados judicialmente. Pode ser visto quando há um uso "off-label" da aplicação. Deste modo, um estudo aprofundado das reais implicações de usos alternativos da IA criada para determinado fim, pode auxiliar a evitar estas distorções.

Ao final do estudo, os autores concluem que reconhecer o contexto e uso pretendido da IA permite identificar e tratar as fontes de danos discriminatórios, tendo em vista que diferentes fontes de danos exigem diferentes intervenções.

Os problemas que surgem com o reconhecimento facial, por exemplo, são possivelmente enquadráveis no viés de representação ou de avaliação, pois muitos dos *big datas* disponíveis publicamente não representam igualmente diferentes grupos, objetos ou geografias. Ter consciência disso permite que o desenvolvedor colete mais amostras de dados para ensinar o modelo.

Ainda, saber que existe e onde ocorre o viés de agregação pode ser útil para evitar que diagnósticos médicos sejam aplicados genericamente a grupos diferentes (ex. homens e mulheres), sem considerar os diferentes funcionamentos de seus corpos e os diferentes tratamentos cabíveis.

Suresh e Guttag também explicam que podem existir múltiplas fontes de vieses em uma mesma IA, eis que escolhas importantes são feitas em todas as etapas da geração de dados e ciclo de vida do ML, de forma que mitigá-los envolve considerar todos os fluxos do processo de construção do modelo e cada decisão tomada. Sugerem, também, soluções apropriadas para contornar os problemas identificados de acordo com suas respectivas necessidades, as quais podem compor as medidas de um programa de *compliance*.³²

³² SURESH, Harini; GUTTAG, John. Understanding potential sources of harm throughout the machine learning life cycle. **MIT Case Studies in Social and Ethical Responsibilities of Computing**, 10 ago. 2021. Disponível em: <https://mit-serc.pubpub.org/pub/potential-sources-of-harm-throughout-the-machine-learning-life-cycle/release/2?from=7166&to=7866>. Acesso em: 13 nov. 2022. *E-book*.

O estudo descrito neste subitem evidencia que há meios técnicos de identificar vieses discriminatórios desferidos por aplicações de IA, os quais são importantes para tentar solucionar e/ou mitigar tais vieses, sendo necessária a implementação de um processo sério desde a sua projeção, que avalie previamente possíveis riscos a fim de aplicar as medidas necessárias para contorná-los.

Percebe-se que a grande maioria dos vieses é gerado a partir de uma decisão humana ou de uma ausência dela, por isso, desde a sua criação, há que se adotar práticas de governança em cada etapa do ciclo de vida de uma IA.

Inioluwa Deborah Raji e Joy Buolamwini publicaram um artigo demonstrando que, após as auditorias públicas realizadas em ferramentas de reconhecimento facial desenvolvidas e utilizadas por grandes corporações – tais como Microsoft, Face++, IBM, Amazon, Kairos –, as taxas de erro apresentadas na identificação de rostos de mulheres, especialmente as de pele negra, reduziram consideravelmente. Isso porque, após serem expostas publicamente, as empresas envolvidas adotaram procedimentos de correção e avaliação das suas ferramentas.³³

O referido artigo aborda como auditorias externas independentes, semelhantes às auditorias contábeis, podem ser eficazes na melhoria de discriminações algorítmicas, demonstrando que muitas vezes faltam incentivos para auditorias e melhorias nos sistemas de IA e que a exposição pública acaba sendo este “incentivo”.

As autoras se inspiraram no procedimento utilizado por *The National Computer Emergency Readiness Team* (CERT) quando identifica riscos à segurança cibernética, que se utiliza de um procedimento chamado *Coordinated Vulnerability Disclosures* (CVD) para informar às empresas, de forma respeitosa e não antagônica, os riscos e as vulnerabilidades de seus sistemas. Assim, as autoras sugerem a utilização de um modelo semelhante para identificação e correção de vieses algorítmicos.

Raji e Buolamwini demonstram que puderam mapear evoluções significativas nas ferramentas de IA disponibilizadas pelas referidas empresas – as quais

³³ RAJI, Inioluwa Deborah; BUOLAMWINI, Joy. Actionable auditing: investigating the impact of publicly naming biased performance results of commercial ai products. In: **Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society**. 2019. p. 429-435.

corrigiram e atualizaram suas aplicações, além de se engajarem em promover um uso ético da IA – a partir de uma auditoria feita na pesquisa de mestrado de Joy no MIT, chamada de *Gender Shades* (a pesquisa que deu origem ao documentário *Coded Bias*, já citada neste trabalho), onde identificou vieses algorítmicos discriminatórios com relação à cor da pele. A pesquisa foi publicada e conduzida com procedimento parecido com o mencionado acima e gerou melhorias consideráveis no desempenho do reconhecimento facial em mulheres de pele negra, segundo o artigo das autoras.³⁴

Portanto, está provado que o argumento de que as IAs que utilizam *machine learning* se tratam de caixas pretas que produzem resultados sem controle humano não se sustenta para que a empresa e seus administradores se isentem do dever de diligenciar para mitigar os danos e discriminações que suas IAs produzirem, eis que foi possível constatar que existem mecanismos de mitigação que, em casos práticos, causaram a melhoria de determinadas discriminações.

Mesmo nos casos em que as aplicações refletiam vieses históricos da sociedade (ex. racismo e machismo), houve possibilidade de redução dos vieses, o que reforça o reconhecimento do dever de se implementar *compliance* algorítmico também no ciclo de vida da IA, como decorrência do dever de diligência.

Agir diferente pode configurar negligência por parte do administrador e da empresa que utiliza IA e, por meio dela, toma decisões discriminatórias sem imediata retratação, pois, evidentemente, não foram asseguradas todas as cautelas necessárias para evitar as discriminações.

4.3 Compliance algorítmico e o dever de diligência

Consoante foi possível verificar acima, é por meio da criação de fluxos mínimos a serem adotados em cada etapa do ciclo de vida da inteligência artificial que opera por aprendizagem de máquina que se torna possível a mitigação da discriminação algorítmica negativa, causadora de danos a terceiros.

É possível verificar que as gigantes da tecnologia já estão aderindo a instrumentos de governança, publicando seus códigos de conduta para

³⁴ RAJI; BUOLAMWINI, *loc. cit.*

desenvolvimento e uso da IA, a exemplo da Google³⁵ e da IBM³⁶ (sendo que esta última companhia inclusive interrompeu suas pesquisas e operações envolvendo *softwares* de reconhecimento facial³⁷, por receio do risco de uso para vigilância da população pelos governos e defendendo a necessidade de debater o uso desta tecnologia, respeitando seus próprios princípios).

Maria Cristine Branco Lindoso refere que o compromisso das empresas em evitar estas discriminações seria um importante ponto de partida para eliminá-las:

O compromisso ético para evitar discriminação de gênero nessas etapas específicas não demanda, apenas, o respeito à lei, já que, por diversas vezes, os sistemas automatizados irão reproduzir realidades discriminatórias. A doutrina mais recente vem falando de um verdadeiro compromisso positivo por parte dos agentes, à luz da ideia de responsabilidade social empresarial, para que eles adotem medidas corretivas da realidade discriminatória nessa etapa de desenho do algoritmo e de data mining, sendo essa uma das melhores formas para mitigar o risco discriminatório.³⁸

Andriei Gutierrez sustenta a implementação de mais transparência e *accountability* com relação às aplicações de IA (destacando ser *accountability* um dos mais importantes princípios) e afirma que, apesar do estado da arte da tecnologia atual – que ainda possui pontos de opacidade – é possível assegurar estes pressupostos éticos em todas as etapas do ciclo da vida da IA.³⁹

Doneda *et al.* também destacam o papel dos quadros éticos corporativos:

³⁵ ARTIFICIAL Intelligence at Google: our principles. In: GOOGLE, 2023. Disponível em: <https://ai.google/principles/>. Acesso em: 18 nov. 2022.

³⁶ ÉTICA DE IA. In: IBM'S FROM ROADBLOCK TO SCALE: the global sprint towards AI study, 2020. Disponível em: <https://www.ibm.com/br-pt/artificial-intelligence/ethics>. Acesso em: 18 nov. 2022.

³⁷ IBM encerra área de reconhecimento facial e pede reforma da polícia. In: FORBES, 09 jun. 2020. Disponível em: <https://forbes.com.br/forbes-tech/2020/06/ibm-encerra-area-de-reconhecimento-facial-e-pede-reforma-da-policia/>. Acesso em: 18 nov. 2022. (Até o presente momento, não foram encontradas notícias ou informações a respeito da retomada da comercialização/prestação destes produtos/serviços).

³⁸ LINDOSO, Maria Cristine Branco. **Discriminação de gênero em processos decisórios automatizados**. Dissertação (Mestrado em Direito) – Programa de Pós-Graduação em direito. Faculdade de Direito, Universidade de Brasília, Brasília, 2019, p. 94.

³⁹ GUTIERREZ, Andriei. É possível confiar em um sistema de Inteligência Artificial? Práticas em torno da melhoria da sua confiança, segurança e evidências de *accountability*. In: FRAZÃO, Ana; MULHOLLAND, Caitlin (Coord.). **Inteligência artificial e Direito: ética, regulação e responsabilidade**. 2. ed. rev. atual. e ampl. São Paulo: Thomson Reuters Brasil, 2020. p. 81-84.

Como explicado acima, a IA pode ser usada para o bem, e para o mal. Isto levanta preocupações agudas, mas tem a virtualidade de trazer benefícios incriveis para os indivíduos e para a sociedade, assim como para resolver - ou pelo menos atenuar - essas mesmas preocupações. Como complemento à lei e à regulamentação, que já aborda e abrange muitas dessas preocupações (CNIL, 2017), as estruturas éticas corporativas podem ser ferramentas importantes para ajudar as empresas a maximizar os benefícios e oportunidades da IA minimizando os seus riscos e preocupações.⁴⁰

De acordo com o dever de diligência analisado, o ciclo de vida da IA deve envolver pelo menos um mapeamento de possíveis riscos (como parte do dever de se informar e de investigar), sobretudo quando estes riscos já foram amplamente divulgados na mídia, até mesmo para um público leigo (como exemplo os documentários do Netflix e notícias).

Assim, considerados os riscos potenciais da utilização de inteligência artificial, é recomendável (e, como visto no decorrer da pesquisa, deve ser compreendido como devido) que as empresas adotem políticas de governança algorítmica, com foco em *accountability*, ou seja, criando códigos de conduta, relatórios de impacto, modelos de prestação de contas transparentes, etc., sobre todos os critérios adotados para criação de determinada IA (documentando os critérios utilizados para programação, os critérios que nortearam a curadoria do *dataset*, etc.), programas de *compliance*, designando seus princípios éticos, responsabilidades, adoção de práticas como AI Safety⁴¹, prevendo graus e análises de riscos e, sempre que necessário, adoção de etapas como testagem, auditoria interna ou externa, monitoramento, correções, etc.

Também se exige do administrador diligente o dever de investigar outros fatos que podem ocasionar possíveis danos à sociedade anônima, tais como condenações judiciais e outras punições pecuniárias ou prejuízos reputacionais, a exemplo do que

⁴⁰ DONEDA, Danilo Cesar Maganhoto *et al.* Considerações iniciais sobre inteligência artificial, ética e autonomia pessoal. **Pensar – Revista de Ciências Jurídicas**, Fortaleza, v. 23, n. 4, p. 1-17, out/dez. 2018. DOI: 10.5020/2317-2150.2018.8257. Disponível em: <https://ojs.unifor.br/rpen/article/view/8257>. Acesso em: 05 out. 2022. p. 13.

⁴¹ É um campo de pesquisa voltado a mitigar possíveis efeitos indesejados relacionados ao uso de inteligência artificial, pesquisando, por exemplo, formas práticas de transformar princípios éticos, como equidade, em algoritmos – entendendo, inclusive, qual a melhor técnica a depender do risco e da função da aplicação.

ocorreu com a empresa Decolar.com.⁴² Deste modo, devem ser tomadas as providências cabíveis para sua mitigação, além do dever de monitorar constantemente o desenvolvimento de uma aplicação que envolva inteligência artificial que tome decisões, especialmente aquelas que possam gerar algum tipo de discriminação.

Bioni e Luciano⁴³ realizam uma interpretação da LGPD e de outras leis gerais de proteção de dados, e se espelham no RGPD europeu e em algumas legislações americanas para buscar nestas leis gerais o princípio da precaução e, com base nele, exigir a implementação de governança e relatórios de impacto no uso de IA, já que esta tecnologia basicamente se alimenta e aprende com base nos dados que recebe, que em sua grande maioria são dados pessoais.

No âmbito do *compliance* anticorrupção, Natalia Cristina Chaves e Marcelo Andrade Féres⁴⁴ avaliam sobre como a IA pode ser útil para efetivação destes controles, mas também reforçam como a adaptação de programas de conformidades, originalmente advindos das leis anticorrupção, podem inspirar um *e-compliance*, conformando o uso da IA às normas aplicáveis.

Ao final do artigo, os autores explicitam o seu conceito de *e-compliance*, que é um *compliance* digital, conformando à IA às normas de proteção do consumidor, dos dados pessoais, dos crimes cibernéticos, sendo uma mistura do *compliance* com direito digital (terminologia). Nas palavras dos autores:

⁴² A empresa Decolar.com foi condenada pelas práticas de discriminação algorítmica geográfica, conhecidas como *geopricing* (discriminar consumidores de determinadas localidades cobrando preços maiores ou diferentes condições do que quando ofertados para consumidores de outras localidades, o que é identificado com base nos dados georreferenciais do dispositivo de pesquisa do serviço/mercadoria) e *geoblocking* (ofertar serviços/mercadorias para consumidores de determinadas localidades e não ofertar os mesmos serviços/mercadorias para consumidores de outras localidades, o que é identificado com base nos dados georreferenciais do dispositivo de pesquisa do serviço/mercadoria). Para mais informações: MARTINS, Guilherme Magalhães. O *geopricing* e *geoblocking* e seus efeitos nas relações de consumo. In: MULHOLLAND, Caitlin; FRAZÃO, Ana (Coord.). **Inteligência artificial e direito: ética, regulação e responsabilidade**. São Paulo: Thomson Reuters Brasil, 2019. p. 651-736.; MEDON, Felipe. **Inteligência artificial e responsabilidade civil: autonomia, riscos e solidariedade**. São Paulo: JusPodivm, 2022. p. 287-290.

⁴³ BIONI, Bruno Ricardo; LUCIANO, Maria. O princípio da precaução na regulação de inteligência artificial: seriam as leis de proteção de dados o seu portal de entrada. In: MULHOLLAND, Caitlin; FRAZÃO, Ana (Coord.). **Inteligência artificial e direito: ética, regulação e responsabilidade**. São Paulo: Thomson Reuters Brasil, 2019. *E-book*.

⁴⁴ CHAVES, Natália Cristina; FÉRES, Marcelo Andrade. Responsabilidade administrativa de pessoas jurídicas na lei anti-corrupção: sanções e critérios de fixação. In: FÉRES, Marcelo Andrade; CHAVES, Natália Cristina. **Sistema anticorrupção e empresa**. Belo Horizonte: D'Plácido, 2018.

Segundo Urs Gasser e Daniel M. Haeusermann, o e-compliance consiste na “gestão de riscos na interseção de lei, tecnologia e mercado que surgiram mediante e em reação à informatização e à rede digital”. De acordo com referidos autores, tais riscos abarcam preocupações com segurança (combate a vírus, spiiwares, worms e hackers, além de roubo de dados e de hardwares); privacidade de dados (tanto de clientes e funcionários); proteção ao consumidor, a qual deverá levar em consideração as legislações dos diversos países em que tal público for alcançado; propriedade intelectual, especialmente direitos autorais; além de governança de conteúdo. Analisando-se esses novos riscos trazidos pela era digital, verificam-se 4 (quatro) características do e-compliance e que o diferenciam do compliance tradicional: “1. Direito e tecnologia digital estão inseparavelmente interconectados; 2. E-compliance tem que lidar e gerenciar riscos legais decorrentes da incerteza jurídica criada pelo ambiente de mercúrio do presente e futuro direito de TI. 3. A digitalização e a expansão da internet levaram a uma ‘internacionalização’ mais pronunciada tanto de questões legais antigas e novas. 4. A dinâmica da digitalização e a incerteza legal associada aumentaram a relevância do soft law.”⁴⁵

Assim, o artigo conclui que o uso da tecnologia pode servir como ferramenta de facilitação do *compliance* e como o *compliance* pode ser utilizado como ferramenta para mitigação de riscos causados pela tecnologia dentro das empresas.

Em semelhante orientação, Frazão ⁴⁶ já em 2019 indicava que independentemente de tratar-se de tecnologia desenvolvida pela própria empresa ou terceirizada, a sua escolha e a sua implementação devem:

- a) ser precedidas do conhecimento dos aspectos fundamentais do *design* e do funcionamento do sistema;
- b) fazer análise de risco;
- c) indagar quais os cuidados foram tomados na programação;

⁴⁵ CHAVES, Natália Cristina; FÉRES, Marcelo Andrade. Responsabilidade administrativa de pessoas jurídicas na lei anti-corrupção: sanções e critérios de fixação. In: FÉRES, Marcelo Andrade; CHAVES, Natália Cristina. **Sistema anticorrupção e empresa**. Belo Horizonte: D'Plácido, 2018.

⁴⁶ FRAZÃO, Ana. Responsabilidade civil dos administradores de sociedades empresárias por decisões tomadas com base em sistemas de inteligência artificial. In: FRAZÃO, Ana; MULHOLLAND, Caitlin (Coord.). **Inteligência artificial e Direito: ética, regulação e responsabilidade**. 2. ed. rev. atual. e ampl., p. 481-525, São Paulo: Thomson Reuters Brasil, 2019. p. 508.

- d) definir quais os *inputs* e *outputs* esperados;
- e) definir quais são as linhas gerais do processamento;
- f) avaliar como os engenheiros ou programadores responsáveis identificaram e endereçaram os principais riscos;
- g) verificar que medidas foram tomadas para evitar resultados equivocados ou discriminatórios;
- h) definir quais dados alimentam o sistema;
- i) verificar as características fundamentais do processamento e suas finalidades;
- j) organizar como será feito o monitoramento e a correção dos resultados indesejados.

Assim, mesmo que após todas as cautelas necessárias o sistema venha acarretar danos, a responsabilidade pessoal do administrador pode ser afastada caso se demonstre que ele foi diligente na escolha e na vigilância da IA utilizada, demonstrando que cumpriu o seu dever de diligência.

Frazão cita que a LGPD também orienta alguns cuidados que devem ser tomados pelos administradores, tendo em vista que determina expressamente algumas medidas diligentes (princípio da transparência, *accountability*, determinações práticas, tais como relatório de impacto etc.). Ressalta-se que a LGPD indica apenas uma parte das medidas pertinentes para o uso de IA, especialmente se esta se utiliza de dados pessoais em seus processos de aprendizagem de máquina.⁴⁷ Contudo, o administrador não pode se restringir apenas a um *compliance* de dados, pois, como visto, esta é apenas uma das diversas fontes de vieses que podem se originar durante o ciclo de vida de uma IA.

A autora reforça, portanto, que adotar uma estrutura compatível é um dever do administrador para ser diligente:

O que importa ressaltar é que o dever de diligência dos administradores de sociedades empresárias apresenta grande potencial para assegurar a adequada

⁴⁷ Apesar da relevância de muitas das medidas impostas pela LGPD, ela está mencionada apenas como exemplo, pois a proteção de dados pessoais não foi objeto desta pesquisa, e por isso a sua respectiva legislação não foi escopo de uma análise mais detida, conforme descrito no texto.

utilização da inteligência artificial na atividade empresarial, inclusive sob o aspecto organizacional. Com efeito, se os administradores têm a obrigação de adotar organização compatível com o risco assumido, é certo que, ao escolherem qualquer sistema de inteligência artificial, precisam saber a real extensão dos riscos a ele inerentes, a fim de que possam controlar adequadamente os referidos riscos.⁴⁸

Com base no Guia que a Comissão Europeia desenvolveu em abril de 2019, intitulado de Diretrizes Éticas para a IA Confiável, Frazão analisa os quatro princípios éticos indicados no guia: respeito pela autonomia humana, prevenção de danos, justiça e explicabilidade.

Convertendo em ações práticas, a autora destaca que os sistemas de IA devem ser centrados no benefício humano desde o *design*, assegurando-se a supervisão humana em todas as etapas do ciclo da vida.

Reforça a importância da não discriminação/estigmatização e o equilíbrio entre interesses conflitantes. Ressalta a necessidade de identificar-se o responsável pelas tomadas de decisão algorítmica, de forma que o processo de tomada de decisão seja explicável à pessoa atingida.

Para isso, os processos demandam transparência e comunicação acessível, resguardando-se a rastreabilidade e auditabilidade do iter decisório algorítmico. O sistema deve ser pensado de modo a possibilitar alguma forma de supervisão humana, bem como a avaliação de *feedback* dos usuários e afetados, de modo a ajustar o que possa ser nocivo a direitos fundamentais.

A autora também ressalta a importância de orientar os usuários sobre o funcionamento do sistema, de modo que seja capaz de compreender os resultados e interagir com a ferramenta.

O guia mencionado pela autora também engloba sete exigências: (i) *human agency* e supervisão humana, (ii) robustez técnica e segurança, (iii) privacidade e governança de dados, (iv) transparência, (v) diversidade, não discriminação e justiça,

⁴⁸ FRAZÃO, Ana. Responsabilidade civil dos administradores de sociedades empresárias por decisões tomadas com base em sistemas de inteligência artificial. In: FRAZÃO, Ana; MULHOLLAND, Caitlin (Coord.). **Inteligência artificial e Direito: ética, regulação e responsabilidade**. 2. ed. rev. atual. e ampl., p. 481-525, São Paulo: Thomson Reuters Brasil, 2019. p. 508.

(vi) bem-estar e ambiental e social e (vii) *accountability*. Além da observância de tais exigências, vale destacar:

Mais do que isso, é importante assegurar que agentes públicos tenham a habilidade de supervisionar tais sistemas por meio de mecanismos que podem ser exigidos em vários graus para apoiar segurança e medidas de controle, dependendo da área de aplicação ou dos riscos de cada sistema de inteligência artificial. A ideia fundamental é que, mantendo-se todas as condições, quanto menos supervisão humana puder haver sobre um sistema de inteligência artificial, mais testes extensivos e uma rígida governança deve haver.⁴⁹

Outro ponto que o guia aborda com grande relevância é a necessidade de documentar cada passo do ciclo de vida da IA, o que vai ao encontro do estudado com relação ao dever de diligência: deve-se documentar os critérios de escolha da base de dados, suas características, planejamento, treinamento, testagem, implementação, etc. Tal medida, além de evidenciar o nível de diligência dedicada pela administração da empresa, é imprescindível para atingimento dos princípios e exigências da IA ética (tais como transparência, explicabilidade, centralidade no ser humano etc.), bem como para mitigar e corrigir danos que venham a ser causados.

Com relação às medidas aptas a evitar discriminação, além das já citadas, menciona-se a diversidade de opiniões e a inclusão de diversos interessados ao longo dos ciclos de vida de IA, de modo a evitar que vieses históricos sejam replicados, o que já deve ser observado desde a coleta dos dados utilizados.

A exigência de *accountability* também guarda relação com o dever de diligência, na medida em que viabiliza auditoria sobre o sistema, a prestação de contas em caso de efeitos indesejados, o que mais uma vez, demonstra a importância de identificar, avaliar e documentar todos os passos do sistema a ser implementado.

⁴⁹ FRAZÃO, Ana. Responsabilidade civil dos administradores de sociedades empresárias por decisões tomadas com base em sistemas de inteligência artificial. In: FRAZÃO, Ana; MULHOLLAND, Caitlin (Coord.). **Inteligência artificial e Direito: ética, regulação e responsabilidade**. 2. ed. rev. atual. e ampl., p. 481-525, São Paulo: Thomson Reuters Brasil, 2019. p. 514.

A autora comenta que o guia reconhece que sempre que não for possível atender a todos os princípios e exigências, este risco deve ser avaliado e mitigado, renunciando-se a dispositivos de IA que ofereçam impactos considerados inaceitáveis. Demonstra que o *compliance* é visto como necessário, sendo imperioso investir em estruturas para assegurar o cumprimento de todas as etapas acima mencionadas.

Atento às melhores práticas, o Fórum Econômico Mundial recentemente publicou recomendações de pelo menos cinco ações a serem adotadas pelas empresas para o desenvolvimento e uso de algoritmos de inteligência artificial responsável, sendo elas:⁵⁰

- k) **Estabelecer princípios de IA:** a empresa deve listar os objetivos legais e éticos que buscará respeitar ao utilizar IA, os quais devem estar alinhados com as necessidades e contexto de cada empresa. Eles serão o norte de todo o ciclo de vida da IA desenvolvida/utilizada. Para tanto, é necessário envolvimento de uma equipe mais diversa do que os desenvolvedores de TI;
- l) **Estabelecer uma estrutura de governança de IA responsável:** sugerem, inicialmente, um mapeamento de todas as áreas afetadas pelo uso de IA, bem como de suas funcionalidades, o que deve ocorrer junto ao setor de tecnologia, de forma interdisciplinar. A partir daí é que serão avaliados os riscos envolvidos e a criação de uma estrutura para melhorar os processos internos, as tecnologias envolvidas, mitigando danos indesejados e modificando as culturas da empresa, com orientações para os desenvolvedores, operadores e usuários finais. Comentam a respeito do número crescente de organizações que podem orientar os esforços da administração, e citam como exemplo a BSA Framework⁵¹, a GAO Risk Framework⁵² e a NIST AI Risk Management

⁵⁰ FIRTH-BUTTERFIELD, Kay; VOGEL, Miriam. 5 ways to avoid artificial intelligence bias with 'responsible AI'. **World Economic Forum**. 05 jul. 2022. Disponível em: <https://www.weforum.org/agenda/2022/07/5-governance-tips-for-responsible-ai/>. Acesso em: 13 nov. 2022.

⁵¹ THE BSA framework for secure software: a new approach to securing the software lifecycle. *In*: BSA THE SOFTWARE ALLIANCE, set. 2020. Disponível em: <https://www.bsa.org/reports/updated-bsa-framework-for-secure-software>. Acesso em: 19 nov. 2022.

⁵² ARTIFICIAL Intelligence: an accountability framework for federal agencies and other entities. *In*: U.S. GOVERNMENT ACCOUNTABILITY OFFICE – GAO. 2021. Disponível em: <https://www.gao.gov/products/gao-21-519sp>. Acesso em: 19 nov. 2022.

Framework⁵³. Descrevem que o EqualAI Framework⁵⁴ fornece cinco pilares para a governança de IA responsável e que o Fórum Econômico Mundial produziu o AI C-Suite Toolkit⁵⁵ para líderes do setor de IA;

m) **Operacionalizar uma estrutura de governança:** após estabelecer qual será a estrutura adequada de governança de acordo com o porte e nível de risco das IAs utilizadas nas empresas, é necessário colocar o projeto em prática, por isso este subitem é dividido em outros quatro: (i) definir um responsável: designar quem serão as pessoas/equipes responsáveis pelo constante acompanhamento e monitoramento das atividades envolvendo IA, seus ajustes e melhorias, bem como pelo atendimento de outras demandas que se tornem necessárias, tais como esclarecimentos, prestação de contas, novos desenvolvimentos etc. O texto destaca a importância dos *feedbacks* e avaliações dos membros da equipe envolvidos na estrutura de governança, bem como dos usuários, operadores e desenvolvedores, para sua constante adaptação, tendo em vista que é a partir do uso em grande escala que os problemas aparecem com maior clareza. Comenta que o time envolvido precisa ser diverso para ter visibilidade dos vários aspectos envolvidos, incluindo especialistas técnicos, engenheiros, especialistas em ética, designers, advogados, desenvolvedores de produtos, executivos de vendas e relações com o consumidor. Relata que "*Diversity makes AI safer as it envisions the broadest range of risks. It also makes it better, as it is crafted for a very broad consumer base.*"⁵⁶; (ii) Comunique: para assegurar que a IA não produza resultados indesejados, é necessário testá-la antes de colocá-la em prática de forma definitiva. Deste modo, é importante que se estabeleçam as etapas do desenvolvimento, os resultados esperados, os riscos (resultados que não devem ocorrer), o que e como precisa ser resolvido, para que isso seja divulgado aos membros da empresa, permitindo que todos monitorem e auxiliem com

⁵³ AI RISK management framework. *In*: NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY – NIST. Set. 2022. Disponível em: <https://www.nist.gov/itl/ai-risk-management-framework>. Acesso em: 19 nov. 2022.

⁵⁴ OUR MISSION is to work with companies, policy makers and experts to reduce bias in our AI. *In*: EQUAL AI, [2022?]. Disponível em: <https://www.equalai.org/about-us/mission/>. Acesso em: 19 nov. 2022.

⁵⁵ EMPOWERING AI leadership: AI C-Suite toolkit. *In*: WORLD ECONOMIC FORUM, jan. 2022. Disponível em: <https://www.weforum.org/reports/empowering-ai-leadership-ai-c-suite-toolkit/>. Acesso em: 19 nov. 2022.

⁵⁶ Em tradução livre: a diversidade torna a IA mais segura, pois prevê a mais ampla gama de riscos. Também a torna melhor, pois é criado para uma base de consumidores muito ampla.

feedbacks. (iii) Documentar estes processos:⁵⁷ o fórum ainda recomenda que estas etapas, descobertas, falhas, riscos, resultados sejam documentados, de forma a dar mais transparência ao processo, evidenciando as cautelas tomadas até para mitigar possíveis responsabilizações. Indicam que a documentação precisa mencionar características dos dados utilizados para treinamento, os resultados, as taxas de sucesso e de insucesso. (iv) Realizar auditorias rotineiras: sugerem a realização de auditorias periódicas, não só nas IAs novas, mas nas já utilizadas;⁵⁸

n) **Treinamento:** fornecer cursos sobre ética em inteligência artificial, explicar os princípios perseguidos pela empresa, quais os deveres e obrigações e quais pontos devem ser atentados. Sugerem, ainda, que altos executivos participem de cursos e certificações de órgãos externos e reconhecidos para aprender melhores práticas.

o) **Questionários/checklists:** itens básicos que precisam ser avaliados/respondidos sempre que uma nova IA estiver sendo projetada/desenvolvida/monitorada/ auditada. Se muitos itens necessários para evitar danos não estiverem supridos, a ferramenta precisará passar por avaliações e ajustes, sendo inclusive submetida a um comitê de ética da empresa.

Ainda vale mencionar que há outros atores internacionais como IEEE⁵⁹, a ISO⁶⁰ e a Partnership on AI⁶¹ engajados em elaborar guias, *standards* e orientações de cunho prático, para empresas e desenvolvedores, indicando quais ações podem ser tomadas para evitar vieses e outros riscos.

⁵⁷ Tal item, além de ter sido indicado por Lupion e Konrad como um dos deveres essenciais para se comprovar o cumprimento do dever de diligência, também é útil para processos futuros, eis que evidencia os caminhos que deram certo e os que deram errado, evitando a repetição de erros.

⁵⁸ Tendo em vista que estas se adaptam para repetir os novos dados que recebem, é necessário auditar constantemente as IAs que já estão em uso. Tais auditorias permitirão que se façam ajustes sempre que necessário.

⁵⁹ ADVANCING Technology for Humanity. In: INSTITUTO DE ENGENHEIROS ELETRICISTAS E ELETRÔNICOS – IEEE, 2023. Disponível em: <https://www.ieee.org/>. Acesso em: 19 nov. 2022.

⁶⁰ ISO/IEC 23053:2022 - Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML). In: ISO, jun. 2022. Disponível em: <https://www.iso.org/standard/74438.html> Acesso em: 19 nov. 2022.

⁶¹ PARTNERSHIP on AI is bringing together diverse voices from across the AI community. In: PARTNERSHIP ON AI, 2023. Disponível em: <https://partnershiponai.org>. Acesso em: 19 nov. 2022.

No Brasil, a entidade Transparência Brasil emitiu uma cartilha sobre governança específica para uso da inteligência artificial no poder público, com indicações a serem respeitadas por órgãos de todas as esferas, que podem ser adaptadas e servir como um guia para auxiliar também as empresas e seus responsáveis pela IA.⁶²

Além disso, a versão criada por juristas brasileiros para criar o Marco Legal da Inteligência Artificial também traz diversas ferramentas e orientações que podem ser observadas pelas empresas enquanto não houver sua aprovação definitiva (e impositiva). Além de traçar metas e objetivos contra a discriminação algorítmica, a redação propõe uma categorização dos riscos conforme o tipo e aplicação de cada IA, classificando as potencialmente discriminatórias como de risco excessivo (quando oferece prejuízo à saúde ou à segurança de grupos vulnerabilizados ou quando utilizada pelo poder público para fins de *social score*⁶³ – art. 14, II e III) ou alto risco (art. 17).

A proposta veda o uso de IAs que ofereçam risco excessivo e impõe as seguintes medidas para IAs de alto risco, com ênfase em vieses discriminatórios:

Art. 20. Além das medidas indicadas no art. 19, os agentes de inteligência artificial que forneçam ou operem sistemas de alto risco adotarão as seguintes medidas de governança e processos internos:

- I – documentação, no formato adequado ao processo de desenvolvimento e à tecnologia usada, a respeito do funcionamento do sistema e das decisões envolvidas em sua construção, implementação e uso, considerando todas as etapas relevantes no ciclo de vida do sistema, tais como estágio de design, de desenvolvimento, de avaliação, de operação e de descontinuação do sistema;
- II – uso de ferramentas de registro automático da operação do sistema, de modo a permitir a avaliação de sua acurácia e robustez e a apurar potenciais

⁶² TRANSPARÊNCIA BRASIL. **Recomendações de governança** - uso de inteligência artificial pelo Poder Público. Fev. 2020. Disponível em: https://www.transparencia.org.br/downloads/publicacoes/Recomendacoes_Governanca_Uso_IA_PoderPublico.pdf. Acesso em: 18 nov. 2022.

⁶³ O uso do *social score* é popularmente atribuído ao governo chinês para controle de sua população, e foi definido pelo projeto de lei brasileiro como: [são vedadas a implementação e uso] pelo poder público, para avaliar, classificar ou ranquear as pessoas naturais, com base no seu comportamento social ou em atributos da sua personalidade, por meio de pontuação universal, para o acesso a bens e serviços e políticas públicas, de forma ilegítima ou desproporcional.

discriminatórios, bem como a implementação das medidas de mitigação de riscos adotadas, com especial atenção para efeitos adversos;

III – realização de testes para avaliação de níveis apropriados de confiabilidade, conforme o setor e o tipo de aplicação do sistema de inteligência artificial, incluindo testes de robustez, acurácia, precisão e cobertura;

IV – medidas de gestão de dados para mitigar e prevenir vieses discriminatórios, incluindo:

a) avaliação dos dados com medidas apropriadas de controle de vieses cognitivos humanos que possam afetar a coleta e organização dos dados, bem como medidas para evitar a geração de vieses por problemas na classificação, falhas ou falta de informação em relação a grupos afetados, falta de cobertura ou distorções em representatividade, conforme a aplicação pretendida, bem como medidas corretivas para evitar a incorporação de vieses sociais estruturais que possam ser perpetuados e ampliados pela tecnologia;

b) composição de equipe inclusiva responsável pela concepção e desenvolvimento do sistema, orientada pela busca da diversidade.

V – adoção de medidas técnicas para viabilizar a explicabilidade dos resultados dos sistemas de inteligência artificial e de medidas para disponibilizar aos operadores e potenciais impactados informações gerais sobre o funcionamento do modelo de inteligência artificial empregado, explicitando a lógica e os critérios relevantes para a produção de resultados, bem como, mediante requisição do interessado, disponibilizar informações adequadas que permitam a interpretação dos resultados concretamente produzidos, respeitado o sigilo industrial e comercial.

Parágrafo único. A supervisão humana de sistemas de inteligência artificial de alto risco buscará prevenir ou minimizar os riscos para direitos e liberdades das pessoas que possam decorrer de seu uso normal ou de seu uso em condições de utilização indevida razoavelmente previsíveis, viabilizando que as pessoas responsáveis pela supervisão humana possam:

I – compreender as capacidades e limitações do sistema de inteligência artificial e controlar devidamente o seu funcionamento, de modo que sinais de anomalias, disfuncionalidades e desempenho inesperado possam ser identificados e resolvidos o mais rapidamente possível;

II – ter ciência da possível tendência para confiar automaticamente ou confiar excessivamente no resultado produzido pelo sistema de inteligência artificial;

III – interpretar corretamente o resultado do sistema de inteligência artificial tendo em conta as características do sistema e as ferramentas e os métodos de interpretação disponíveis;

IV – decidir, em qualquer situação específica, por não usar o sistema de inteligência artificial de alto risco ou ignorar, anular ou reverter seu resultado; e

V – intervir no funcionamento do sistema de inteligência artificial de alto risco ou interromper seu funcionamento.⁶⁴

A avaliação de impacto algorítmico proposta também traça um importante processo metodológico para assegurar as cautelas necessárias ou uso de IA:

Art. 24. A metodologia da avaliação de impacto conterà, ao menos, as seguintes etapas:

I – preparação;

II – cognição do risco;

III – mitigação dos riscos encontrados;

IV – monitoramento.

§ 1º A avaliação de impacto considerará e registrará, ao menos:

a) riscos conhecidos e previsíveis associados ao sistema de inteligência artificial à época em que foi desenvolvido, bem como os riscos que podem razoavelmente dele se esperar;

b) benefícios associados ao sistema de inteligência artificial;

c) probabilidade de consequências adversas, incluindo o número de pessoas potencialmente impactadas;

d) gravidade das consequências adversas, incluindo o esforço necessário para mitigá-las;

e) lógica de funcionamento do sistema de inteligência artificial;

f) processo e resultado de testes e avaliações e medidas de mitigação realizadas para verificação de possíveis impactos a direitos, com especial destaque para potenciais impactos discriminatórios;

⁶⁴ CUEVA, Ricardo Villas Bôas *et al.* **Relatório final da Comissão de Juristas responsável por subsidiar elaboração de substitutivo sobre Inteligência Artificial no Brasil**. Brasília: Coordenação de Comissões Especiais, Temporárias e Parlamentares de Inquérito, 2022. Disponível em: <https://legis.senado.leg.br/sdleg-getter/documento?dm=9221643&ts=1671480646036>. Acesso em: 04 jan. 2023.

g) treinamento e ações de conscientização dos riscos associados ao sistema de inteligência artificial;

h) medidas de mitigação e indicação e justificação do risco residual do sistema de inteligência artificial, acompanhado de testes de controle de qualidade frequentes;

i) medidas de transparência ao público, especialmente aos potenciais usuários do sistema, a respeito dos riscos residuais, principalmente quando envolver alto grau de nocividade ou periculosidade à saúde ou segurança dos usuários, nos termos dos artigos 9º e 10 da Lei nº 8.078, de 11 de setembro de 1990 (Código de Defesa do Consumidor);

§ 2º Em atenção ao princípio da precaução, quando da utilização de sistemas de inteligência artificial que possam gerar impactos irreversíveis ou de difícil reversão, a avaliação de impacto algorítmico levará em consideração também as evidências incipientes, incompletas ou especulativas.

[...]

Art. 25. A avaliação de impacto algorítmico consistirá em processo iterativo contínuo, executado ao longo de todo o ciclo de vida dos sistemas de inteligência artificial de alto risco, requeridas atualizações periódicas.⁶⁵

Também vale lembrar que apesar de tais esforços demandarem muito tempo, atenção e custos, pesquisas já demonstraram que a adoção de boas práticas que podem ser utilizadas por desenvolvedores para mitigação de vieses discriminatórios, ou seja, um *compliance* algorítmico, gera resultados efetivos, não configurando apenas um registro formal de que se buscou agir com cautela. Tais afirmações puderam ser evidenciadas nas pesquisas de Raji e Buolamwini (a qual demonstrou que todas as empresas denunciadas tiveram melhor desempenho em suas ferramentas de reconhecimento facial ao revisarem seus sistemas adotando melhores práticas) e do estudo de Suresh e Gutttag (que demonstrou como identificar

⁶⁵ CUEVA, Ricardo Villas Bôas *et al.* **Relatório final da Comissão de Juristas responsável por subsidiar elaboração de substitutivo sobre Inteligência Artificial no Brasil**. Brasília: Coordenação de Comissões Especiais, Temporárias e Parlamentares de Inquérito, 2022. Disponível em: <https://legis.senado.leg.br/sdleg-getter/documento?dm=9221643&ts=1671480646036>. Acesso em: 04 jan. 2023.

as diferentes fontes de vieses e sugeriu soluções tecnicamente aplicáveis para a sua mitigação).

Sendo assim, consoante verificou-se por meio de doutrina jurídica e de duas pesquisas da área técnica computacional, o *compliance* é visto como uma ferramenta eficaz na mitigação de discriminações algorítmicas quando observado durante o ciclo de vida da inteligência artificial que se utiliza de aprendizagem de máquina, além de ser uma forma eficiente de documentar/demonstrar o cumprimento do dever de diligência.

4.4 Conclusões parciais

No presente capítulo, foi possível compreender que o uso de inteligência artificial está cada vez mais presente nas rotinas pessoais e, especialmente, nas rotinas empresariais, eis que estas, além de se utilizarem das facilidades oferecidas pelas ferramentas, são as responsáveis por ofertarem elas ao público.

Assim, ao decidir delegar atividades empresariais a aplicações de IA, as empresas e os seus administradores têm o dever de zelar pelos resultados que a IA produz, tomando todas as cautelas possíveis para evitar e/ou ao menos mitigar os eventuais efeitos colaterais que ocasionar a terceiros. Deste modo, há necessidade de respeitar o dever de diligência sempre que se optar por delegar alguma atividade empresarial a uma IA.

Por conseguinte, analisou-se como o dever de diligência pode ser concretizado e reconhecido nestas situações, constatando-se que os instrumentos adequados para sua instrumentalização seriam a criação de estruturas de governança, especialmente o *compliance* – o que, por via reflexa, auxilia no cumprimento de outro dever: o de legalidade, ou seja, o dever de não desrespeitar normas legais.

Também foi possível verificar, por meio da doutrina examinada, que a governança e o *compliance* (inicialmente desenvolvidos para outras finalidades) possuem etapas e ferramentas adaptáveis para diversas atividades e podem ser desenvolvidos para avaliação e prevenção de potenciais danos decorrentes de aplicações de inteligência artificial. Não bastasse isso, constatou-se que tais

métodos são algumas das importantes medidas que sugerem os cientistas da área de tecnologia que estudam como desenvolver e usar inteligência artificial de forma ética e segura, os quais podem e devem ser fonte de consulta para os administradores diligentes, fortes no *right to rely*.

Ao final do capítulo, abordou-se o que a doutrina entende por melhores práticas corporativas a serem implementadas para demonstrar a diligência dos administradores que se valem da inteligência artificial em suas rotinas empresariais, concluindo-se pela existência de diversas medidas de cunho prático que podem e devem consolidadas em um programa de *compliance* específico, com o intuito de evitar ou ao menos mitigar discriminações algorítmicas negativas em cumprimento ao dever de diligência.

5. CONCLUSÃO FINAL

Como se viu no decorrer da pesquisa, os avanços tecnológicos influenciam diretamente no mundo jurídico, pois criam deveres decorrentes das novas formas de (des)cumprir normas, as quais podem surgir a partir da mudança dos meios de se executar tarefas. Além disso, os avanços também possibilitam novos recursos que subsidiam o cumprimento de deveres, adaptando a forma de cumpri-los e até mesmo o sentido que pode ser dado às obrigações legais.

Quando a presente investigação iniciou, o objetivo era compreender, a partir da interpretação doutrinária, se havia algum dever de cuidado que pudesse ser exigido das empresas e de seus gestores no uso de inteligência artificial, assim como já era amplamente reconhecido para outras decisões negociais, com base no art. 153 da Lei 6.404/76.

Isso porque, pesquisas recentes como as das professoras Lais Machado Lucas e Ana Frazão, relatadas no texto, demonstravam que o conteúdo do dever de diligência deve ser interpretado à luz do contexto em que a empresa e o administrador estão inseridos.

Tais pesquisas possuíam objetos diversos ao que foi aqui analisado: a da professora Lais Machado Lucas interpretou o conteúdo do dever de diligência para demonstrar o dever de implementar um *compliance* anticorrupção, enquanto a da professora Ana Frazão analisou o dever de diligência dos administradores que utilizam inteligência artificial para justificar a responsabilidade civil destes. Maria Cristina Borrasca Felisberto também analisou ambos os temas, porém, com foco na *Business Judgment Rule*.

Apesar dos escopos serem distintos, as conclusões das pesquisas acima serviram de inspiração para a investigação ora realizada, despertando o interesse por averiguar o conteúdo atual do dever de diligência à luz do estado da arte da inteligência artificial e da discriminação algorítmica, trazendo, portanto, uma nova abordagem sobre este tema.

Constatou-se, então, que as evoluções das ferramentas computacionais vêm refletindo diretamente na interpretação que se dá ao dever de diligência: a tecnologia,

especialmente a inteligência artificial, trouxe novos focos de atenção para os administradores de empresas que lidam com estas aplicações em alguma atividade. Em paralelo, a tecnologia também contribuiu para a disseminação de informações e possibilita a análise de um volume maior de dados, viabilizando formas de se informar, controlar e gerir que não existiam.

Neste sentido, verificou-se que um importante ponto de atenção dos administradores deve ser a inteligência artificial utilizada para tomadas de decisões (ou sugestões) que influenciam na vida de terceiros, pois estas podem ser discriminatórias. Esta possibilidade indesejada já foi evidenciada em diversas pesquisas, sendo amplamente divulgadas em meios como jornais, revistas, documentários, entre outras mídias, o que deve provocar o alerta (*red flags*) para os gestores de empresas que lidam com esta tecnologia, sejam próprias ou terceirizadas.

Deste modo, com a conclusão da presente pesquisa, é possível afirmar que:

a) a inteligência artificial está em constante expansão nas rotinas empresariais e pessoais, sobretudo em virtude da possibilidade de (i) identificar padrões em grandes volumes de dados, (ii) correlacionar esses dados (iii) classificar grupos de características comuns, aplicando tais conclusões de forma estatística em eventos futuros e permitindo que os sistemas "adivinhem" situações ou realizem tarefas com muita facilidade;

b) os benefícios gerados a partir da inteligência artificial surgem acompanhados de determinados efeitos colaterais que podem influenciar negativamente a vida de pessoas e grupos, de forma que isso restringe direitos e garantias individuais, como o direito à igualdade, configurando o que se convém chamar de discriminação algorítmica;

c) a discriminação algorítmica foi categorizada em grupos por diversos autores analisados, a partir de diferentes premissas, tais como suas origens ou seus efeitos e, tais categorizações, apesar de não serem unânimes, evidenciaram que cada etapa do desenvolvimento de um sistema de inteligência artificial, deve exigir um determinado tipo de atenção e de cuidados para que não seja prejudicial e discriminatória a terceiros, especialmente quando utilizam aprendizagem de máquina;

d) os alertas sobre a discriminação algorítmica vêm sendo amplamente divulgados em pesquisas acadêmicas e mídias, de modo que os agentes que utilizam e/ou ofertam inteligência artificial devem preocupar-se com a criação de mecanismos que garantam a adoção de cautelas específicas em cada etapa do ciclo de vida da IA, na tentativa de evitar, ou ao menos mitigar, a perpetração de vieses discriminatórios negativos;

e) as cautelas no uso ou desenvolvimento da inteligência artificial podem exigir pesquisa, investimento de tempo e de dinheiro, envolvimento de partes interessadas, reconhecimento (e até exposição) de vulnerabilidades, contratação de terceiros para assessoramento, auditorias e/ou até prejuízos e retrabalho;

f) a decisão de implementar (e de como implementar) inteligência artificial precisa ser encarada como uma decisão estratégica que pode influenciar diretamente nos negócios empresariais, exigindo que as orientações e a disseminação da cultura de cuidados partam da própria gestão;

g) o administrador diligente não pode sujeitar a empresa que administra a correr os riscos legais, financeiros e reputacionais que podem decorrer de algum resultado discriminatório sem antes empregar máximos cuidados e diligências, como exige a previsão do art. 153 da Lei 6.404/76 – cuidados estes que devem ser aplicáveis também às empresas que usam ou desenvolvem inteligência artificial;

h) a interpretação do que significa cumprir o dever de diligência se adaptou conforme (i) o passar do tempo, (ii) a evolução das legislações, (iii) as práticas do mercado, (iv) as informações que se tornaram disponíveis a partir do aprimoramento técnico, do desenvolvimento de novas profissões e especialidades, entre outras variáveis;

i) a subjetividade da previsão legal da Lei das Sociedades Anônimas permite que em cada caso concreto se analise seu cumprimento. Inclusive, nas pesquisas doutrinárias mais atuais, a menção a um dever de *compliance* e de boas práticas de governança são vistas como aquelas condutas que concretizam e formalizam muitos dos subdeveres relacionados ao dever de diligência, sendo apontado como um novo subdever;

j) os precedentes da Comissão de Valores Mobiliários analisados reforçaram o entendimento doutrinário, demonstrando que o dever de diligência

deve ser interpretado à luz da época dos fatos e das práticas de mercado usuais do momento em que a decisão empresarial foi tomada. Evidenciaram, ainda, que situações fáticas provocam novas demandas que alteram a forma de cumprimento do dever de diligência, como criação de estruturas de *compliance*, que passou a ser uma exigência do mercado e não mais uma opção, sobretudo tudo após a Operação Lava-Jato, que revelou a importância destas práticas empresariais;

k) havendo uma nova realidade fática decorrente da disseminação do uso de IA, o dever de diligência deve ser respeitado pelos administradores das empresas sempre que se decidir delegar atividades empresariais a aplicações de IA, zelando-se de forma eficiente pelos resultados que a IA produz, tomando todas as cautelas possíveis para evitar e/ou ao menos mitigar os eventuais efeitos colaterais que podem ocasionar a terceiros;

l) o dever de diligência será cumprido quando o administrador criar estruturas de governança, especialmente um programa de *compliance* algorítmico – o que, por via reflexa, auxilia no cumprimento de outro dever: o de legalidade, ou seja, o dever de não desprezar normas legais;

m) a governança e o *compliance* são as ferramentas indicadas para cumprimento dos deveres legais, bem como representam importantes medidas para o desenvolvimento e para o uso da inteligência artificial de forma ética e segura, como demonstram a doutrina, o Fórum Econômico Mundial e o projeto de lei do Marco Legal da Inteligência Artificial.

Então, respondendo aos problemas norteadores da investigação realizada, foi possível concluir que o ordenamento jurídico brasileiro espera que os administradores das empresas que desenvolvem ou utilizam inteligência artificial cumpram um dever de *compliance* e de boas práticas de governança algorítmica, revelando-se, portanto, um novo subdever que concretiza e formaliza muitos dos outros subdeveres identificados na composição do dever de diligência, tais como: se informar, se qualificar, bem administrar, investigar, vigiar, supervisionar/monitorar, intervir, não praticar erros graves, participar, entre outros.

Sendo assim, exige-se dos administradores a implementação de um *compliance* algorítmico eficiente que considere todas as etapas do ciclo de vida da inteligência artificial e que trace estratégias para cada um dos desafios identificados

no decorrer do seu desenvolvimento, pois estas ações representam a diligência esperada de um administrador que deseja cumprir os deveres estabelecidos no art. 153 da Lei das SA.

Deste modo, a presente pesquisa contribuiu para o cenário acadêmico, jurídico e empresarial ao constatar um novo subdever derivado do dever de diligência, que é o dever de estruturação de um *compliance* algorítmico, que mapeie os riscos da IA a ser desenvolvida e/ou utilizada, com atenção para os possíveis focos originários de discriminação. Ato seguinte, a pesquisa também contribuiu indicando quais as medidas mínimas que devem ser realizadas e documentadas ao constatar as origens de uma potencial discriminação, com foco em sua mitigação ou correção.

Verificou-se, assim, que os administradores de empresas que desejam desenvolver ou utilizar IA em qualquer tarefa devem, portanto, acautelar-se sobre todos os impactos que a atividade que será realizada pela máquina vai gerar na vida dos seus usuários ou partes afetadas, a fim de que uma potencial discriminação negativa seja evitada ou, caso surja posteriormente de forma involuntária, seja sanada.

Neste sentido, entende-se que para que o desenvolvimento ou utilização de uma ferramenta de inteligência artificial ocorra de forma diligente, o administrador deve previamente:

Reunir informações e documentos sobre a inteligência artificial que pretende ser utilizada, bem como do tipo de técnica que servirá para o seu aprimoramento (regras pré-definidas, *machine learning*, entre outras), mapeando todo seu ciclo e informando-se, ainda, a respeito dos bancos de dados utilizados para o seu treinamento (origem, conteúdo, rótulos etc.), quando cabível. Estas informações e documentos precisam ser qualificados, ou seja, oriundos de fontes seguras e confiáveis (mídias reconhecidamente confiáveis, pesquisas acadêmicas, livros, documentários, empresa responsável).

Depois, é necessário investigar, ou seja, examinar criteriosamente as informações obtidas, buscando compreender aspectos legais, éticos, práticos, financeiros, sociais etc., da IA a ser utilizada, bem como do seu respectivo banco de dados, devendo contar com apoio de especialistas quando não dispor de conhecimento técnico para tanto, nos termos do *right to rely*. Se, nesta análise, o

administrador observar a existência de potenciais danos (*red flags*), deve avaliá-los com mais profundidade, verificando alternativas e soluções para contornar os obstáculos ou, então, justificar plausivelmente porque o risco é válido para a empresa (desde que não acarrete descumprimentos legais ou danos intencionais).

Além disso, mesmo após a decisão tomada, o administrador deve monitorar o seu objeto constantemente. Ou seja, supervisionar (diretamente ou por intermédio de terceiros delegados) o andamento da IA implementada para acompanhar as consequências e os resultados, recalculando rotas e adotando melhorias em tempo hábil sempre que se perceber a necessidade de algum ajuste na decisão tomada. E sem esquecer de manter registros e evidências idôneos de todas estas etapas, quase como um prontuário. Tudo isso deve estar no foco de um programa de *compliance* algorítmico, sendo sua estrutura e dimensão adequadas às peculiaridades e recursos disponíveis para cada empresa.

Portanto, ao passo que (i) o dever de diligência se atualiza conforme o estado da arte em que as decisões negociais estão inseridas, havendo reconhecimento de diversos novos subdeveres desde a publicação da Lei 6.404/76, interpretando-se o *compliance* como um dos mais recentes subdeveres, e que, (ii) no atual estado da arte, a inteligência artificial é uma ferramenta com potencial de causar danos decorrentes de discriminação, não há como negar que as estruturas de *compliance* e governança algorítmicos são mecanismos aptos para evitar ou mitigar estes efeitos colaterais indesejados, tornando-se, então, um novo dever a ser cumprido pelos administradores que desejam ser reconhecidos como cumpridores do dever de diligência.

REFERÊNCIAS

ADVANCING Technology for Humanity. *In*: INSTITUTO DE ENGENHEIROS ELETRICISTAS E ELETRÔNICOS – IEEE, 2023. Disponível em: <https://www.ieee.org/>. Acesso em: 19 nov. 2022.

AI RISK management framework. *In*: NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY – NIST. Set. 2022. Disponível em: <https://www.nist.gov/itl/ai-risk-management-framework>. Acesso em: 19 nov. 2022.

ANGWIN, Julia *et al.* Machine Bias: there's software used across the country to predict future criminals. And it's biased against blacks. **ProPublica**, 23 maio 2016. Disponível em: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. Acesso em: 15 mar. 2022.

ARTIFICIAL Intelligence at Google: our principles. *In*: GOOGLE, 2023. Disponível em: <https://ai.google/principles/>. Acesso em: 18 nov. 2022.

ARTIFICIAL INTELLIGENCE INDEX – AI; STANFORD INSTITUTE FOR HUMAN-CENTERED ARTIFICIAL INTELLIGENCE. **The AI Index Report, 2021**. Disponível em: <https://aiindex.stanford.edu/report/#individual-chapters>. Disponível em: 10 out. 2022.

ARTIFICIAL Intelligence: an accountability framework for federal agencies and other entities. *In*: U.S. GOVERNMENT ACCOUNTABILITY OFFICE – GAO. 2021. Disponível em: <https://www.gao.gov/products/gao-21-519sp>. Acesso em: 19 nov. 2022.

ASSOCIAÇÃO PARA PROMOÇÃO DA EXCELÊNCIA DO SOFTWARE BRASILEIRO – SOFTEX. **Observatório Softex 2019**. Overview do setor de tecnologia da informação brasileiro nos últimos dez anos. Brasília, 2019. *E-book*. Disponível em: <https://softex.br/download/overview-do-setor-de-tecnologia-da-informacao-brasileiro-nos-ultimos-dez-anos/?wpdmdl=97749&masterkey=5cee9abb82a39>. Acesso em: 17 out. 2022.

BIDDLE, Sam. The internet's new favorite AI proposes torturing Iranians and surveilling mosques. **The Intercept**. 08 dez. 2022. Disponível em: <https://theintercept.com/2022/12/08/openai-chatgpt-ai-bias-ethics/>. Acesso em: 09 jan. 2023.

BIONI, Bruno Ricardo; LUCIANO, Maria. O princípio da precaução na regulação de inteligência artificial: seriam as leis de proteção de dados o seu portal de entrada. *In*: MULHOLLAND, Caitlin; FRAZÃO, Ana (Coord.). **Inteligência artificial e direito: ética, regulação e responsabilidade**. São Paulo: Thomson Reuters Brasil, 2019. *E-book*.

BRASIL. Comissão de Valores Mobiliários. **Processo Administrativo Sancionador CVM nº 05/2016**. Processo Eletrônico SEI nº 19957.010647/2019-97. Voto Diretor Gustavo Machado Gonzales. Fls. 45 e 46. 2020. Disponível em: https://www.gov.br/cvm/pt-br/assuntos/noticias/anexos/2020/20201103_PAS_CVM_SEI_19957_010647_2019_97_05_2016_voto_diretor_gustavo_gonzalez.pdf-4796a160193148b2a835fb6458616c1c. Acesso em: 17 out. 2022.

BRASIL. Comissão de Valores Mobiliários. **Processo Administrativo Sancionador CVM nº RJ2014/8013**. Voto Diretor Relator Gustavo Machado Gonzales. Fls. 45 e 46. 2018. Disponível em: https://conteudo.cvm.gov.br/sancionadores/sancionador/2018/20180731_PAS_RJ20148013.htm. Acesso em: 09 jan. 2023.

BRASIL. **Lei 10.406, de 10 de janeiro de 2002**. Institui o Código Civil. Brasília: Presidência da República, 2002. Disponível em: https://www.planalto.gov.br/ccivil_03/LEIS/2002/L10406compilada.htm. Acesso em: 10 out. 2022.

BRASIL. **Lei n. 6.404, de 15 de dezembro de 1976**. Dispõe sobre as sociedades por ações. Brasília: Casa Civil, 1976. Disponível em: https://www.planalto.gov.br/ccivil_03/Leis/L6404consol.htm. Acesso em: 15 abr. 2022.

BRIGAGÃO, Pedro Henrique Castello. **A administração das companhias e a business judgment rule: uma análise à luz do direito brasileiro**. 191 p. Monografia (Graduação em Direito). Pontifícia Universidade Católica do Rio de Janeiro: Rio de Janeiro, 2014.

BRIGAGÃO, Pedro Henrique Castello. **Dever de diligência dos administradores de sociedades anônimas**. Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Direito. Rio de Janeiro, 2012. Disponível em: https://www.puc-rio.br/ensinopesq/ccpg/pibic/relatorio_resumo2012/resumos_pdf/ccs/DIR/Pedro%20Henrique%20Castello%20Brigag%E3o.pdf. Acesso em: 17 out. 2022.

BUOLAMWINI, Joy; GEBRU, Tinmit. "Gender shades: intersectional accuracy disparities in commercial gender classification." **Proceedings of Machine Learning Research**, v. 81, n. 1, p. 1–15, 2018. Conference on Fairness, Accountability, and Transparency.

CALLEJÓN, Francisco Balaguer. La constitución del algoritmo. **Fundación Manuel Giménez Abad de Estudios Parlamentarios y del Estado Autonomico**. Zaragoza, 2022.

CAMPOS, Luiz Antonio de Sampaio. Deveres e responsabilidades. In: LAMY FILHO, Alfredo; PEDREIRA, José Luiz Bulhões (Coord.). **Direito das companhias**. São Paulo: Forense, 2009.

CHAVES, Natália Cristina; FÉRES, Marcelo Andrade. Responsabilidade administrativa de pessoas jurídicas na lei anti-corrupção: sanções e critérios de fixação. *In*: FÉRES, Marcelo Andrade; CHAVES, Natália Cristina. **Sistema anticorrupção e empresa**. Belo Horizonte: D'Plácido, 2018.

CODED BIAS. *In*: WIKIPÉDIA: a enciclopédia livre. [San Francisco: Wikimedia Foundation, 2021]. Disponível em: https://en.wikipedia.org/wiki/Coded_Bias. Acesso em: 23 jan. 2023.

COELHO, Fábio Ulhoa (Coord.) *et al.* **Lei das sociedades anônimas comentada**. Rio de Janeiro: Forense, 2021.

CORRÊA, Nicholas Kluge. **Artificial Intelligence ethics and safety**: practical tools for creating "good" models. 2021. p. 4. DOI: 10.13140/RG.2.2.16827.13602. Disponível em: https://www.researchgate.net/publication/357032443_Artificial_Intelligence_Ethics_and_Safety_practical_tools_for_creating_good_models. Acesso em: 13 nov. 2022.

CORRÊA, Nicholas Kluge. **Ética e segurança da inteligência artificial**: ferramentas práticas para se criar "bons" modelos. Dez. 2021. Disponível em: https://www.researchgate.net/publication/357032340_Etica_e_Seguranca_da_Inteligencia_Artificial_ferramentas_praticas_para_se_criar_bons_modelos. Acesso em: 13 nov. 2022.

CORRÊA, Nicholas Kluge *et al.* Worldwide AI ethics: a review of 200 guidelines and recommendations for AI governance. **Veritas**, Porto Alegre, 2022. Disponível em: <http://dx.doi.org/10.48550/arXiv.2206.11922>. Acesso em: 14 jan. 2023.

CUEVA, Ricardo Villas Bôas *et al.* **Relatório final da Comissão de Juristas responsável por subsidiar elaboração de substitutivo sobre Inteligência Artificial no Brasil**. Brasília: Coordenação de Comissões Especiais, Temporárias e Parlamentares de Inquérito, 2022. Disponível em: <https://legis.senado.leg.br/sdleg-getter/documento?dm=9221643&ts=1671480646036>. Acesso em: 04 jan. 2023.

CUEVA, Ricardo Villas Bôas. Funções e finalidades dos programas de compliance. *In*: CUEVA, Ricardo Villas Bôas; FRAZÃO, Ana (Coord.). **Compliance**: perspectivas e desafios dos programas de conformidade. Belo Horizonte: Fórum, 2018.

CVM conclui julgamento de processos envolvendo Petrobras. *In*: COMISSÃO DE VALORES MOBILIÁRIOS. Brasília, 03 nov. 2020. Disponível em: <https://www.gov.br/cvm/pt-br/assuntos/noticias/cvm-conclui-julgamento-de-processos-envolvendo-petrobras-2756185ec313490fae197ce3735d6927>. Acesso em: 17 out. 2022.

DIAS, Natasha. Panorama da adoção de tecnologia por empresas brasileiras. **Cortex**, 30 ago. 2022. Disponível em: <https://www.cortex->

intelligence.com/intelligence-review/panorama-sobre-adocao-de-tecnologia-por-empresas-brasileiras. Acesso em: 17 out. 2022.

DOMINGOS, Pedro. **O algoritmo mestre**: como a busca pelo algoritmo de machine learning definitivo recriará nosso mundo. 1ª edição. São Paulo: Novatec, 2017.

DONEDA, Danilo Cesar Maganhoto *et al.* Considerações iniciais sobre inteligência artificial, ética e autonomia pessoal. **Pensar – Revista de Ciências Jurídicas**, Fortaleza, v. 23, n. 4, p. 1-17, out/dez. 2018. DOI: 10.5020/2317-2150.2018.8257. Disponível em: <https://ojs.unifor.br/rpen/article/view/8257>. Acesso em: 05 out. 2022.

DUARTE, Alan; NEGÓCIO, Ramon de Vasconcelos. Todos são iguais perante o algoritmo? uma resposta cultural do Direito à discriminação algorítmica. **Revista Direito Público**, Brasília, v. 18, n. 100, 2022. DOI: 10.11117/rdp.v18i100.5869. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/5869>. Acesso em: 4 ago. 2022.

DZIEZA, Josh. Why captchas have gotten so difficult: demonstrating you're not a robot is getting harder and harder. **The Verge**. 01 fev. 2019. Disponível em: <https://www.theverge.com/2019/2/1/18205610/google-captcha-ai-robot-human-difficult-artificial-intelligence>. Acesso em: 15 mar. 2022.

EIZIRIK, Nelson. A lei das S/A comentada. 2. ed. rev. e ampl. Artigos 138-205. São Paulo: Quartir Latin, 2015. v. II.

EMPOWERING AI leadership: AI C-Suite toolkit. *In*: WORLD ECONOMIC FORUM, jan. 2022. Disponível em: <https://www.weforum.org/reports/empowering-ai-leadership-ai-c-suite-toolkit/>. Acesso em: 19 nov. 2022.

ÉTICA DE IA. *In*: IBM'S FROM ROADBLOCK TO SCALE: the global sprint towards AI study, 2020. Disponível em: <https://www.ibm.com/br-pt/artificial-intelligence/ethics>. Acesso em: 18 nov. 2022.

FACCHINI NETO, Eugênio; SCALZILLI, Roberta. Pode a ética controlar o desenvolvimento tecnológico? O caso da inteligência artificial, à luz do direito comparado. *In*: COLOMBO, Cristiano. FALEIROS JR, José Luiz de Moura; ENGELMANN, Wilson. **Tutela jurídica do corpo eletrônico**: novos desafios ao direito digital. Indaiatuba/SP: Foco, 2022. *E-book*.

FELISBERTO, Maria Cristina Borrasca. Business judgment rule: o limite dos riscos assumidos pelos administradores e o uso da Inteligência Artificial. **Revista Brasileira de Direito Comercial**, [S.l.], n. 43, out./nov. 2021.

FERRARI, Isabela *et al.* **Justiça digital**. 1. ed. São Paulo: Thomsom Reuters Brasil, 2020.

FIRTH-BUTTERFIELD, Kay; VOGEL, Miriam. 5 ways to avoid artificial intelligence bias with 'responsible AI'. **World Economic Forum**. 05 jul. 2022. Disponível em: <https://www.weforum.org/agenda/2022/07/5-governance-tips-for-responsible-ai/>. Acesso em: 13 nov. 2022.

FÓRUM Stack Overflow. **Developer survey**. 2021. Disponível em: <https://insights.stackoverflow.com/survey/2021#developer-profile-demographics>. Acesso em: 15 set. 2022.

FRAZÃO, Ana. Discriminação algorítmica: compreendendo o que são os julgamentos algorítmicos e o seu alcance na atualidade. Parte I. **JOTA**. 16 jun. 2021. Disponível em: <https://www.jota.info/opiniao-e-analise/colunas/constituicao-empresa-e-mercado/discriminacao-algoritmica-16062021>. Acesso em: 18 nov. 2021.

FRAZÃO, Ana. Responsabilidade civil dos administradores de sociedades empresárias por decisões tomadas com base em sistemas de inteligência artificial. *In*: FRAZÃO, Ana; MULHOLLAND, Caitlin (Coord.). **Inteligência artificial e Direito: ética, regulação e responsabilidade**. 2. ed. rev. atual. e ampl. São Paulo: Thomson Reuters Brasil, 2020. *E-book*.

FRAZÃO, Ana. Responsabilidade civil dos administradores de sociedades empresárias por decisões tomadas com base em sistemas de inteligência artificial. *In*: FRAZÃO, Ana; MULHOLLAND, Caitlin (Coord.). **Inteligência artificial e Direito: ética, regulação e responsabilidade**. 2. ed. rev. atual. e ampl., p. 481-525, São Paulo: Thomson Reuters Brasil, 2019.

FRAZÃO, Ana. Dever de diligência: Novas perspectivas em face de programas de compliance e de atingimento de metas. **JOTA**. 15 fev. 2017. Disponível em: <http://jota.info/colunas/constituicao-empresa-e-mercado/dever-de-diligencia-15022017>. Acesso em: 17 out. 2022.

FRAZÃO, Ana; CUEVA, Ricardo Villas Bôas (Org.). **Compliance: perspectivas e desafios dos programas de conformidade**. 1. ed. Belo Horizonte: Forum, 2018. v. 1. 527p.

FRAZÃO, Ana; MULHOLLAND, Caitlin (Coord.). **Inteligência artificial e direito: ética, regulação e responsabilidade**. 2. ed. ver. atual. e ampl. São Paulo: Thomson Reuters Brasil, 2020. *E-book*.

FREITAS, Cinthia Obladen de Almendra; BARDDAL, Jean Paul. Análise preditiva e decisões judiciais: controvérsia ou realidade? **Revista Democracia Digital e Governo Eletrônico**, Florianópolis, v. 1, n. 18, p. 107-126, 2019.

FREITAS, Juarez; FREITAS, Thomas Bellini. **Direito e Inteligência Artificial: em defesa do humano**. 1. ed. Belo Horizonte: Forum, 2020.

GABRIEL, Martha. **Inteligência artificial: do zero ao metaverso**. 1. ed. Barueri/SP: Atlas, 2022.

GERALDO RUFINO: o ex-catador de latinhas que criou uma das maiores empresas de reciclagem de caminhões. *In*: INFOMONEY, 2023. Disponível em: <https://www.infomoney.com.br/perfil/geraldo-rufino/>. Acesso em: 16 out. 2022.

GHOSH, Iman. 4 key areas where AI and IoT are being combined. **World Economic Forum**, 15 mar. 2021. Disponível em: <https://www.weforum.org/agenda/2021/03/ai-is-fusing-with-the-internet-of-things-to-create-new-technology-innovations/>. Acesso em: 12 nov. 2022.

GONZALES, Gustavo Machado; CORRÊA, Bruno Tostes. Dever de legalidade dos administradores de sociedades. **Revista Semestral de Direito Empresarial**, n. 30 (janeiro/junho 2022), Rio de Janeiro: Renovar, 2007.

GUTIERREZ, Andriei. É possível confiar em um sistema de Inteligência Artificial? Práticas em torno da melhoria da sua confiança, segurança e evidências de accountability. *In*: FRAZÃO, Ana; MULHOLLAND, Caitlin (Coord.). **Inteligência artificial e Direito: ética, regulação e responsabilidade**. 2. ed. rev. atual. e ampl. São Paulo: Thomson Reuters Brasil, 2020.

HARARI, Yuval Noah. **Homo Deus: uma breve história do amanhã**. São Paulo: Companhia das Letras, 2016.

HOFFMANN-RIEM, Wolfgang. Inteligência artificial como oportunidade de regulação jurídica. **Revista de Direito Público**, Brasília, n. 90, v. 16, p. 11-38, nov./dez. 2019.

IBM encerra área de reconhecimento facial e pede reforma da polícia. *In*: FORBES, 09 jun. 2020. Disponível em: <https://forbes.com.br/forbes-tech/2020/06/ibm-encerra-area-de-reconhecimento-facial-e-pede-reforma-da-policia/>. Acesso em: 18 nov. 2022.

ISO/IEC 23053:2022 - Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML). *In*: ISO, jun. 2022. Disponível em: <https://www.iso.org/standard/74438.html> Acesso em: 19 nov. 2022.

KAUFMAN, Dora. **A inteligência artificial irá suplantar a inteligência humana? [S.l.]**: Estação das Letras e Cores Edi, 2019. *E-book*.

KONRAD, Carlos Bender; LUPION, Ricardo. Dever de diligência dos administradores em tempos de pandemia e impactos da crise nos contratos empresariais. **RJLA, [S.l.]**, ano 7, n. 3, p. 427-456, 2021.

KUNICHOFF, Yana; SIER, Patrick. The contradictions of Chicago police's secretive list. **Chicago Magazine**, 21 ago. 2017. Disponível em:

<https://www.chicagogmag.com/city-life/august-2017/chicago-police-strategic-subject-list/>. Acesso em: 15 mar. 2022.

LAVADO, Thiago. Aumento do uso de reconhecimento facial pelo poder público no Brasil levanta debate sobre limites da tecnologia. **G1**, 21 fev. 2020. Disponível em: <https://g1.globo.com/economia/tecnologia/noticia/2020/02/21/aumento-do-uso-de-reconhecimento-facial-pelo-poder-publico-no-brasil-levanta-debate-sobre-limites-da-tecnologia.ghtml>. Acesso em: 15 mar. 2022.

LEE, Kai-Fu. **Inteligência artificial**: como os robôs estão mudando o mundo, a forma como amamos, nos relacionamos, trabalhamos e vivemos. Tradução: Marcelo Barbão. 1. ed. Rio de Janeiro: Globo Livros, 2019. *E-book*.

LEE, Peter. Learning from Tay's introduction. **Microsoft**, 25 mar. 2016. Disponível em: <https://blogs.microsoft.com/blog/2016/03/25/learning-tays-introduction/#sm.00000gjdppwcfcus11t6oo6dw79gw>. Acesso em: 18 nov. 2022.

LINDOSO, Maria Cristine. **Discriminação de gênero em processos decisórios automatizados**. Dissertação (Mestrado em Direito) – Programa de Pós-Graduação em direito. Faculdade de Direito, Universidade de Brasília, Brasília, 2019.

LUCAS, Laís Machado. **Programas de integridade nas sociedades anônimas**: implementação como conteúdo do dever de diligência dos administradores. Porto Alegre: Livraria do Advogados, 2021.

LUPION, Ricardo. Contratos empresariais e o mundo BANI: Brittle, Anxious, Nonlinear and Incomprehensible. **Revista Jurídica Luso-Brasileira - RJLB, [S.l.]**, ano 8, n. 6, p. 1942-1944, 2022.

MAGRANI, Eduardo; SILVA, Priscilla; VIOLA, Rafael. Novas perspectivas sobre ética e responsabilidade de inteligência artificial – Parte I – Compreensão da inteligência artificial e dos seus pressupostos de controle e regulação. *In*: FRAZÃO, Ana; MULHOLLAND, Caitlin. **Inteligência artificial e Direito**. São Paulo: Revista dos Tribunais. 2020. *E-book*.

MARTINS, Guilherme Magalhães. O geoprising e geoblocking e seus efeitos nas relações de consumo. *In*: MULHOLLAND, Caitlin; FRAZÃO, Ana (Coord.). **Inteligência artificial e direito**: ética, regulação e responsabilidade. São Paulo: Thomson Reuters Brasil, 2019. p. 651-736.; MEDON, Felipe. **Inteligência artificial e responsabilidade civil**: autonomia, riscos e solidariedade. São Paulo: JusPodivm, 2022.

MENDES, Laura Schertel; MATTIUZZO, Marcela. Discriminação algorítmica: conceito, fundamento legal e tipologia. **Revista Direito Público**, Brasília, v. 16, n. 90, 2019. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/3766>. Acesso em: 6 ago. 2022.

MIOTTO, Rafael. Significados de 'patroa' e 'mulher-solteira' mudam no Google depois de críticas de Anitta e Luísa Sonza. **G1**, 18 set. 2020. Disponível em: <https://g1.globo.com/economia/tecnologia/noticia/2020/09/18/significados-de-patroa-e-mulher-solteira-mudam-no-google-depois-de-criticas-de-anitta-e-luisa-sonza.ghtml> Acesso em: 15 mar. 2022.

MORRE fundador da marca Azaleia Calçados. *In*: AGÊNCIA ESTADO. Folha de Londrina. 24 jan. 2004. Disponível em: <https://www.folhadelondrina.com.br/geral/morre-fundador-da-marca-azaleia-de-calcados-477027.html> Acesso em: 17 out. 2022.

NUNES, Dierle; MARQUES, Ana Luiza. Inteligência artificial e direito processual: vieses algorítmicos e os riscos de atribuição de função decisória às máquinas. **Revista de Processo**, [S.l.], v. 285, p. 421-447, nov. 2018.

O'NEIL, Cathy. **Algoritmos de destruição em massa**: como o big data aumenta a desigualdade e ameaça à democracia. Tradução Rafael Abraham. 1. ed. Santo André/SP: Rua do Sabão, 2020.

OLIVA, Milena Donato; SILVA, Rodrigo da Guia. Origem e evolução histórica do compliance no Direito brasileiro. *In*: FRAZÃO, Ana; CUEVA, Ricardo Villas Bôas (Org.). **Compliance**: perspectivas e desafios dos programas de conformidade. 1. ed. Belo Horizonte: Forum, 2018. v. 1. 527p.

OLIVEIRA, Daniele de Lima. **Deveres e responsabilidade dos administradores da S/A**. Dissertação (Mestrado em Direito) - Faculdade de Direito da Pontifícia Universidade Católica de São Paulo, São Paulo, 2008.

OPENAI. [Site institucional]. 2023. Disponível em: <https://openai.com/about/>. Acesso em: 09 jan. 2023.

OUR MISSION is to work with companies, policy makers and experts to reduce bias in our AI. *In*: EQUAL AI, [2022?]. Disponível em: <https://www.equalai.org/about-us/mission/>. Acesso em: 19 nov. 2022.

PALMA, Amanda; PACHECO, Clarissa. Entenda como funciona o reconhecimento facial que ajudou a prender mais de 100 na BA - Câmeras estão espalhadas em várias partes de Salvador e imagens são enviadas à central da SSP. **Correio 24 horas**. Salvador, 05 jan. 2020. Disponível em: <https://www.correio24horas.com.br/noticia/nid/entenda-como-funciona-o-reconhecimento-facial-que-ajudou-a-prender-mais-de-100-na-ba>. Acesso em: 15 mar. 2022.

PARENTE, Flávia. **O dever de diligência dos administradores de sociedades anônimas**. Rio de Janeiro: Renovar, 2005.

PARGENDLER, Mariana. Responsabilidade civil dos administradores e business judgment rule no Direito Brasileiro. **Revista dos Tribunais**, São Paulo, v. 953, p. 51-74, 2015.

PARTNERSHIP on AI is bringing together diverse voices from across the AI community. *In: PARTNERSHIP ON AI*, 2023. Disponível em: <https://partnershiponai.org>. Acesso em: 19 nov. 2022.

PICCOLI, Maria Eduarda Ferreira; GONÇALVES, Oksandro Osdival. A responsabilidade civil dos administradores no Brasil e a Business Judgment Rule. **Revista Jurídica Luso-Brasileira - RJLB**, [S.l.], ano 8, n. 6, p. 1497-1530, 2022.

PROCESSOS sancionadores julgados. *In: COMISSÃO DE VALORES MOBILIÁRIOS*. Brasília, 2022. Disponível em: <https://conteudo.cvm.gov.br/sancionadores>. Acesso em: 17 out. 2022.

QUATTRINI, Larissa Teixeira. **Os deveres dos administradores de sociedades anônimas abertas**: estudo de casos. São Paulo: Saraiva, 2014. *E-book*.
RAJI, Inioluwa Deborah; BUOLAMWINI, Joy. Actionable auditing: investigating the impact of publicly naming biased performance results of commercial ai products. *In: Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*. 2019.

RIBEIRO, Renato Ventura. **Dever de diligência dos administradores de sociedades**. São Paulo: Quartier Latin, 2006.

RUSSELL, Stuart.; NORVIG, Peter. **Artificial intelligence**: a modern approach. 4. ed. London: Pearson Education, 2021.

SARLET, Ingo Wolfgang. **Dignidade (da pessoa) humana e direitos fundamentais na Constituição Federal de 1988**. 10. ed. rev. atual. e ampl. Porto Alegre: Livraria do Advogado, 2019.

SARLET, Ingo Wolfgang; SARLET, Gabrielle. Igualdade como proibição de discriminação e direito à (e dever de) inclusão: o acesso ao ensino superior e a regulamentação do estatuto brasileiro das pessoas com deficiência. **Revista Direito Público**, Brasília, v. 14, n. 78, 2017, 197-226 nov./dez. 2017.

SCHWAB, Klaus. **A quarta revolução industrial**. Tradução Daniel Moreira Miranda. São Paulo: Edipro, 2019.

SCHWAB, Klaus. The fourth industrial revolution: what it means, how to respond. **World Economic Forum**, 14 jan. 2016. Disponível em: <https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/>. Acesso em: 12 nov. 2022.

SILVA, Alexandre Couto. **Responsabilidade dos administradores de S/A**: business judgment rule. Rio de Janeiro: Elsevier, 2017.

SILVEIRA, Paulo Antônio Caliendo Velloso da. **Ética e Inteligência Artificial: da possibilidade filosófica de agentes morais artificiais**. 1. ed. Porto Alegre: Fi, 2021. v. 1.

SOARES, Marcelo Negri; KAUFFMAN, Marcos Eduardo; CHAO, Kuo-Ming. Inteligência artificial: impactos no Direito e na Advocacia. **Revista Direito Público**, Brasília, v. 17, n. 93, 2020. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/3555>. Acesso em: 13 ago. 2022.

SOUZA, Ludmilla. Indústria de software e serviços de TIC cresceu 6,5% em 2021. **Agência Brasil**. 19 jul. 2022. Disponível em: <https://agenciabrasil.ebc.com.br/economia/noticia/2022-07/industria-de-software-e-servicos-de-tic-cresceram-65-em-2021>. Acesso em: 17 out. 2022.

STATE v. Loomis: Wisconsin Supreme Court requires warning before use of algorithmic risk assessments in sentencing. **130 Harvard Law Review**, 1530. Mar, 2017. Disponível em: <https://harvardlawreview.org/2017/03/state-v-loomis/>. Acesso em: 15 mar. 2022.

SURESH, Harini; GUTTAG, John. Understanding potential sources of harm throughout the machine learning life cycle. **MIT Case Studies in Social and Ethical Responsibilities of Computing**, 10 ago. 2021. Disponível em: <https://mit-serc.pubpub.org/pub/potential-sources-of-harm-throughout-the-machine-learning-life-cycle/release/2?from=7166&to=7866>. Acesso em: 13 nov. 2022. *E-book*.

THE BSA framework for secure software: a new approach to securing the software lifecycle. *In*: BSA THE SOFTWARE ALLIANCE, set. 2020. Disponível em: <https://www.bsa.org/reports/updated-bsa-framework-for-secure-software>. Acesso em: 19 nov. 2022.

TRANSPARÊNCIA BRASIL. **Recomendações de governança - uso de inteligência artificial pelo Poder Público**. Fev. 2020. Disponível em: https://www.transparencia.org.br/downloads/publicacoes/Recomendacoes_Governanca_Uso_IA_PoderPublico.pdf. Acesso em: 18 nov. 2022.

VECCHIO, Fabrício Bon; VIEIRA, Débora Manke. O controle societário, a teoria Business Judgment Rule e a influência da Inteligência Artificial. *In*: VEIGA, Fábio da Silva; ZALUCKI, Mariusz (Org.). **Legaltech, Artificial Intelligence and the future of legal practice**. Porto/Kraków: Instituto Iberoamericano de Estudos Jurídicos; AFM Kraków University, 2022.

VINCENT, James. Google wants you to help train its AI by labeling images in Google Photos: a new optional feature in Google Photos on Android. **The Verge**. 11 nov. 2020. Disponível em: <https://www.theverge.com/2020/11/11/21559930/google-train-ai-photos-image-labelling-app-android-update>. Acesso em: 15 mar. 2022.

WEBB, Amy. **The Big Nine**: how the tech titans and their thinking machines could warp humanity. New York: PublicAffairs, 2019.

WIMMER, Miriam. Responsabilidade de agentes empresariais por ilícitos administrativos praticados por sistemas de inteligência artificial. *In*: FRAZÃO, Ana; MULHOLLAND, Caitlin (Coord.). **Inteligência artificial e direito**: ética, regulação e responsabilidade. 2. ed. ver. atual. e ampl. São Paulo: Thomsom Reuters Brasil, 2020. *E-book*.

YAZBEK, Otavio. Representação do dever de diligência na doutrina jurídica brasileira: um exercício e alguns desafios. *In*: KUYVEN, Luiz Fernando Martins (Coord.). **Temas essenciais de direito empresarial**: estudos em homenagem a Modesto Carvalhosa, São Paulo: Saraiva, 2012. *E-book*.

