

VIÉS RACIAL EM SISTEMAS DE INTELIGÊNCIA ARTIFICIAL: DESAFIOS PARA A RESPONSABILIDADE CIVIL E A NECESSIDADE DE PARAMETRIZAÇÃO DO RISCO

RACIAL BIAS IN ARTIFICIAL INTELLIGENCE SYSTEMS: CHALLENGES FOR CIVIL LIABILITY AND THE NEED FOR RISK PARAMETERIZATION

Gustavo Silveira Borges *
José Luiz de Moura Faleiros Júnior **

RESUMO: Este artigo aborda os desafios para a responsabilização civil por ilícitos relacionados a falhas de sistemas de inteligência artificial enviesados em relação a questões raciais. Explora-se a possibilidade de um algoritmo ser considerado racista – o que não se confirma nas conclusões – e como isso pode causar danos pela ação omissiva de quem não realiza adequada curadoria de dados que conduzem ao viés algorítmico. Além disso, discute-se a necessidade de parametrização normativa baseada na gradação do risco, em linha com o modelo que segue em discussão na União Europeia. A pesquisa utiliza o método dedutivo, com embasamento bibliográfico e revisão de conceitos teóricos. O objetivo é buscar uma conclusão diante desse tema problemático, considerando a preservação dos direitos humanos e os desafios apresentados pela evolução dos algoritmos e dos sistemas de inteligência artificial.

Palavras-chave: algoritmos; inteligência artificial; viés racial; responsabilidade civil; risco.

ABSTRACT: This article addresses the challenges of civil liability in cases involving wrongdoings related to biased artificial intelligence systems concerning racial matters. It explores the possibility of an algorithm being deemed racist—which is not confirmed in the conclusions—and how this can lead to harm due to the omission of proper data curation, resulting in algorithmic bias. Furthermore, the need for normative parameterization based on risk gradation, in line with the ongoing discussions in the European Union, is discussed. The research adopts a deductive method, relying on bibliographic support and revisiting theoretical concepts. The objective is to reach a conclusion on this problematic issue, considering the preservation of human rights and the challenges presented by the evolution of algorithms and artificial intelligence systems.

Keywords: algorithms; artificial intelligence; racial bias; civil liability; risk.

* Doutor em Direito pela Universidade Federal do Rio Grande do Sul - UFRGS (2013 - CAPES 5), tendo realizado Pós-Doutorado em Direito pela Universidade do Vale do Rio dos Sinos - UNISINOS (2014 - CAPES 6), com bolsa de pesquisa PNPd/CAPES. Mestre em Ciências Criminais pela Pontifícia Universidade Católica do Rio Grande do Sul - PUCRS (2007 - CAPES 5). Especialista em Ciências Penais pela PUCRS). Pós-graduado pela Escola da Magistratura do Rio Grande do Sul - AJURIS. Graduado em Ciências Jurídicas e Sociais pela PUCRS. Atualmente, é Professor da Graduação e Pesquisador permanente do Programa de Pós-Graduação em Direito na Universidade do Extremo Sul Catarinense (PPGD/UNESC). É Coordenador e professor da Especialização em Direito Civil e Processo Civil da UNESC. Pesquisador permanente do Programa de Pós-Graduação em Direito (PPGD/UNESC), vinculado à Linha de Pesquisa Direitos Humanos, Cidadania e Novos Direitos e à Área de Concentração Direitos Humanos e Sociedade. Membro associado do CONPEDI (Conselho Nacional de Pesquisa e Pós-Graduação em Direito). Membro do Instituto Brasileiro de Estudos de Responsabilidade Civil - IBERC. E-mail: gustavoborges@hotmail.com / ORCID: <https://orcid.org/0000-0001-9673-4321>

** Doutorando em Direito Civil pela Faculdade de Direito da Universidade de São Paulo - USP/Largo de São Francisco. Doutorando em Direito, na área de estudo Direito, Tecnologia e Inovação, pela Universidade Federal de Minas Gerais - UFMG. Mestre e Bacharel em Direito pela Universidade Federal de Uberlândia - UFU. Especialista em Direito Digital. Advogado e professor. E-mail: jfaleiros@usp.br / ORCID: <https://orcid.org/0000-0002-0192-2336>

SUMÁRIO: 1. Introdução; 2. Notas sobre o funcionamento de sistemas de inteligência artificial; 3. Algumas sinalizações regulatórias: responsabilidade civil e danos por falhas de algoritmos; 4. Viés racial e responsabilidade civil a partir de falhas algorítmicas: possíveis soluções; 5. Conclusão. Referências.

1. INTRODUÇÃO

Sistemas de inteligência artificial têm se tornado estruturas complexas e cada vez mais avançadas, cuja aplicação se concretiza em verdadeira plêiade de finalidades que produzem resultados otimizados – praticamente instantâneos – e baseados em fórmulas e equações viabilizadas pela heurística computacional.

Tais estruturas evoluem a partir de técnicas como o aprendizado de máquina (*machine learning*), que, aliado à popularização das Tecnologias de Informação e Comunicação (TICs), ao avanço das redes e da telecomunicação, à transformação digital e à virtualização das relações cotidianas, propiciam novo paradigma, com impactos jurídicos sobre toda a ciência. As previsões, embora ainda sejam baseadas na matemática, avançam rumo à superação constante de barreiras e limites que abrem margem a questionamentos que sempre instigaram a curiosidade humana na busca por explicações definitivas sobre os limites da inteligência.

Embora documentos importantíssimos, como o *Artificial Intelligence Act* europeu de 2021 (2021 EU AIA) e o recentíssimo *Algorithmic Accountability Act* norte-americano de 2022 (2022 US AAA), que atualizou a versão anterior, de 2019, busquem se afastar da intrincada discussão terminológica e de debates sobre o alcance semântico do termo “inteligência”, preferindo se reportar a “sistemas decisionais automatizados” (*Automated Decision Systems*, ou ADS’s), a questão que se lança neste estudo é a seguinte: a partir do processamento de dados sobre raça, pode um sistema de IA agir de forma discriminatória e, em última análise, causar danos pela prática de uma “discriminação algorítmica”? Se afirmativa a resposta, como se lhe poderia imputar responsabilidade? Em caso negativo, quem deve ser responsabilizado pelo viés racial?

Sabe-se que o conceito de singularidade tecnológica surge, nesse exato contexto, como fundamento teórico necessário para a compreensão do presente e do futuro de algoritmos complexos que, justamente por se aprimorarem autonomamente, poderão extrapolar os limites da heurística e, eventualmente, se tornar indistinguíveis do modo humano de “pensar”.

Há quem se reporte à nomenclatura *Artificial General Intelligence (AGI)* para descrever o mesmo (e hipotético) fenômeno transformador, especialmente a partir da pujança de técnicas mais avançadas, como o aprendizado profundo (*deep learning*), que fomenta sistemas ditos “generativos”, porquanto baseados em redes neurais recorrentes (*recurrent neural networks*, ou RNNs) ou em *transformers* (modelos de linguagem de larga escala com aplicação no processamento de linguagem natural), o que torna essa realidade mais instigante, especialmente na análise de suas consequências, que são muitas – algumas empolgantes e desafiadoras – ao mesmo tempo que seus riscos representam grandes percalços para a Ciência do Direito.

Fato é que, para o recorte mais específico desse estudo, como se verá, a inquietante indagação lançada como mote não surge desatrelada da práxis, pois notícias recentes já dão

conta de “discriminações algorítmicas” decorrentes de vieses. Resta saber se a adjetivação em comento é plausível e se pode ser imputada a um algoritmo em caso de dano no qual o ilícito, a despeito dos pormenores técnicos, se aproxime da prática discriminatória e repudiável do racismo, quando decorrente de enviesamento.

Trabalhar-se-á com a hipótese de que a preservação dos direitos humanos, mesmo no contexto de incerteza que ainda paira sobre o ritmo da evolução dos algoritmos, é uma realidade atual, impondo-se releituras de sua funcionalização, mesmo enquanto não atingida a mencionada singularidade tecnológica ou a *Artificial General Intelligence (AGI)*.

Em matéria de responsabilidade civil, rechaçando-se a ideia de que nexos de imputação podem ser atribuídos a esses algoritmos, destacar-se-á a necessidade de realce a parâmetros normativos baseados na teoria do risco para que se possa avançar, do ponto de vista dogmático, no sentido da efetivação de um modelo coerente e suficientemente flexível de responsabilização que implique adequada gradação de risco para sistemas de inteligência artificial que possam se tornar enviesados por questões de raça, a depender de análise contextual e necessariamente avaliada segundo gradação condizente com o potencial de enviesamento, tal como se mostra a tendência regulatória da União Europeia (com a abordagem baseada em riscos, ou *risk-based approach*), de modo a inadmitir soluções simplistas como o recurso à teoria da culpa.

A pesquisa será levada a efeito pelo método dedutivo, com suporte bibliográfico doutrinário e revisitações teórico-conceituais para, ao final, explorar-se uma possível conclusão ao tema-problema.

2. NOTAS SOBRE O FUNCIONAMENTO DE SISTEMAS DE INTELIGÊNCIA ARTIFICIAL

Heurística computacional é a expressão que sintetiza o processo cognitivo empregado nas decisões não racionais, definidas a partir de formulações baseadas em aproximações progressivas do resultado almejado¹. Em outras palavras, são fórmulas matemáticas que apresentam soluções preditivas de natureza estatística, indicando cenários com maior ou menor propensão à ocorrência prática a partir de inferências.

Com formulações que remontam à matemática euclidiana², proposições variadas surgiram ao longo dos séculos, evoluindo em verdadeira sintonia com o próprio desenvolvimento científico. Naturalmente, marcos históricos como os experimentos de Thomas Edison com circuitos heteródinos³ ou o emprego de relés eletromecânicos no desenvolvimento de calculadoras como a “Máquina de Chales Babbage”, ou a “Mark I”, de Howard Aiken, são

¹ MICHALEWICZ, Zbigniew; FOGEL, David B. *How to solve it: Modern heuristics*. Cham: Springer, 2000, p. 55.

² GOETTENAUER, Carlos Eduardo. Algoritmos, inteligência artificial, mercados. Desafios ao arcabouço jurídico. In: FRAZÃO, Ana; CARVALHO, Angelo Gamba Prata de (coord.). *Empresa, mercado e tecnologia*. Belo Horizonte: Fórum, 2019, p. 272.

³ ARMSTRONG, Edwin H. Operating field of the audion. *Electrical World*, Nova York, v. 54, n. 1, p. 1149-1152, jul./dez. 1914, p. 1152. Disponível em: <https://bit.ly/2NLe0Qf>. Acesso em: 18 mar. 2023.

exemplos relativamente recentes⁴ – datados da primeira metade do século XX – que ilustram bem como a busca por conforto e praticidade na substituição do labor humano pelo de equipamentos acabou sendo fundamental para o desenvolvimento da eletrônica⁵.

O que não se pode negar é que a mudança de paradigma⁶ vislumbrada desde o início deste novo período tem um substrato essencial: a informação. Tudo muda com Alan Turing, em 1938, e o famoso “Teste de Turing” (*Entscheidungsproblem*), que buscava investigar o potencial de uma máquina para processar informações a ponto de gerar respostas da mesma forma que um humano o faria⁷. Em síntese, esperava-se que o processamento imbatível dos microprocessadores permitisse à máquina, eventualmente, se “emancipar” e, de fato, emular o comportamento humano. O que Turing não esperava era que os conceitos exatos e herméticos da matemática seriam incapazes de, no teste, permitir às máquinas ludibriar, mentir e dissimular, o que tornava fácil a detecção de respostas humanas em comparação às das máquinas, quando realizado o famoso teste. Não obstante, sistemas contemporâneos de inteligência artificial, ditos “generativos”, com destaque para o *Generative Pre-Trained Transformer (GPT)*, já se demonstraram capazes de superar o “Teste de Turing” (e muitos outros testes similares, como o Winograd Schema⁸) e até mesmo exames realizados por humanos, a exemplo de exames para o exercício de profissões como a de médico ou de advogado, nos Estados Unidos da América⁹.

Se a informática marcou um novo estágio de poder computacional e de desenvolvimento de *hardware*, o acúmulo informacional foi o responsável por ‘alimentar’ esses novos equipamentos. Nesse contexto, a solução para o *Entscheidungsproblem* passaria, necessariamente, pelo enfrentamento das principais objeções à proposta de que máquinas podem ‘pensar’. Mais do que um teste, era preciso que se tornasse viável o que Turing batizou

⁴ VALLAS, Steven Peter. *Work*. Cambridge: Polity, 2012, p. 39.

⁵ CERUZZI, Paul E. *A history of modern computing*. 2. ed. Cambridge: The MIT Press, 2003, p. 18-25.

⁶ Thomas Kuhn empregou a expressão *paradigma* no seguinte sentido: “um paradigma é aquilo que os membros de uma comunidade partilham e, inversamente, uma comunidade científica consiste em homens que partilham um paradigma”. O termo *paradigma*, nesta pesquisa, ganha conotação mais ampliada, não só no sentido de uma comunidade científica ou de determinada época, mas diz respeito às diversas mutações do pensamento ocidental. KUHN, Thomas. *A estrutura das revoluções científicas*. São Paulo: Perspectiva, 2000, p. 221.

⁷ Friedrich Kittler, se reportando ao “Teste de Turing” e ao poder da informação, destaca o seguinte: “Only in Turing’s paper On Computable Numbers with an Application to the Entscheidungsproblem there existed a machine with unbounded resources in space and time, with infinite supply of raw paper and no constraints on computation speed. All physically feasible machines, in contrast, are limited by these parameters in their very code. The inability of Microsoft DOS to tell more than the first eight letters of a file name such as WordPerfect gives just a trivial or obsolete illustration of a problem that has provoked not only the ever-growing incompatibilities between the different generations of eight-bit, sixteen-bit and thirty-two-bit microprocessors, but also a near impossibility of digitizing the body of real numbers formerly known as nature”. KITTLER, Friedrich. There is no software. *CTHEORY.net*. 18 out. 1995. Disponível em: <http://www.ctheory.net/articles.aspx?id=74>. Acesso em: 18 mar. 2023.

⁸ FALEIROS JÚNIOR, José Luiz de Moura. Breves reflexões sobre os impactos jurídicos do algoritmo GPT-3. In: BARBOSA, Mafalda Miranda et al (coord.). *Direito digital e inteligência artificial: diálogos entre Brasil e Europa*. Indaiatuba: Foco, 2021, p. 521-524.

⁹ VARANASI, Lakshmi. AI models like ChatGPT and GPT-4 are acing everything from the bar exam to AP Biology. Here’s a list of difficult exams both AI versions have passed. *Business Insider*, 25 jun. 2023. Disponível em: <https://www.businessinsider.com/list-here-are-the-exams-chatgpt-has-passed-so-far-2023-1>. Acesso em: 26 jun. 2023.

de “jogo da imitação”¹⁰: um cenário no qual determinada máquina se tornasse capaz de enganar um terço de seus interlocutores, fazendo-os acreditar que se trataria de um ser humano¹¹. Estaria tal máquina ‘pensando’?

Esse único questionamento desencadeou diversas teorizações acerca da superação das diferenças entre humanos e máquinas (*human-machine divide*)¹² e dos dilemas de desenvolvimento e evolução da inteligência artificial¹³. É possível dizer que, no atual estado da técnica, ainda não é factível a conclusão de que um algoritmo possa se tornar ‘inteligente’¹⁴. Nos dizeres de Howard Gardner, “pode-se concluir que a habilidade lógico-matemática não é um sistema tão “puro” ou “autônomo” como outros revisados aqui, e talvez deva contar não como uma única inteligência, mas como algum tipo de inteligência supra ou mais geral”¹⁵.

De fato, “a percepção desse cenário foi antecedida em muito pela literatura científica, o imaginário popular e a cultura”¹⁶. Para que essa inteligência dotada de características peculiares e especiais seja atingida, tornando-se uma “superinteligência”¹⁷ apta a projetar as condições iniciais para que se produza uma explosão de inteligência sustentável que eleve as condições da vida humana, em harmonia com o desenvolvimento tecnológico, os caminhos do desenvolvimento algorítmico passarão a demandar grande evolução das técnicas de *machine learning*¹⁸.

Alguns festejam a emergência da Quarta Revolução Industrial¹⁹, que promete tornar a

¹⁰ TURING, Alan M. Computing machinery and intelligence. *Mind*, Oxford, n. 236, p. 433-460, out. 1950, p. 25. Disponível em: <https://dx.doi.org/10.1093/mind/LIX.236.433> Acesso em: 18 mar. 2023.

¹¹ Cf. HARNAD, Stevan. The Annotation Game: on Turing (1950), on computing, machinery and intelligence. In: EPSTEIN, Robert; PETERS, Grace (ed.). *Parsing the Turing Test: philosophical and methodological issues in the quest for the thinking computer*. Cham: Springer, 2008.

¹² WARWICK, Kevin. The disappearing human-machine divide. In: ROMPORTL, Jan; ZACKOVA, Eva; KELEMEN, Jozef (ed.). *Beyond Artificial Intelligence: the disappearing human-machine divide*. Cham: Springer, 2015, p. 9.

¹³ FLASINSKI, Mariusz. *Introduction to Artificial Intelligence*. Cham: Springer, 2016, p. 3-13.

¹⁴ KAPLAN, Jerry. *Humans need not apply: a guide to wealth and work in the Age of Artificial Intelligence*. New Haven: Yale University Press, 2015, p. 3-16.

¹⁵ GARDNER, Howard. *Frames of mind: the theory of multiple intelligences*. Nova York: Basic Books, 2011, p. 168, tradução livre. No original: “(...) one could conclude that logical-mathematical ability is not as “pure” or “autonomous” a system as others reviewed here, and perhaps should count not as a single intelligence but as some kind of supra- or more general intelligence.”

¹⁶ DONEDA, Danilo Cesar Maganhoto; MENDES, Laura Schertel; SOUZA, Carlos Affonso Pereira de; ANDRADE, Norberto Nuno Gomes de. 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, p. 2. Os autores prosseguem: “Mesmo a gigantesca capacidade de cálculo, que hoje vemos como intrínseca aos computadores ao se demonstrar uma realidade concreta, foi de início percebida em contraposição direta às capacidades humanas (que eram, afinal, o padrão de comparação por excelência), abrindo a possibilidade de novas comparações possíveis entre o humano e a máquina, entre criação e criatura.”

¹⁷ Cf. BOSTROM, Nick. *Superintelligence: paths, dangers, strategies*. Oxford: Oxford University Press, 2014.

¹⁸ Como diz Pedro Domingos: “With machine learning, computers write their own programs, so we don’t have to. Wow. *Computers write their own programs*. Now that’s a powerful idea, maybe even a little scary. If computers start to program themselves, how will we control them?” DOMINGOS, Pedro. *The master algorithm: how the quest for the ultimate learning machine will remake our world*. Nova York: Basic Books, 2015 p. 6.

¹⁹ SCHWAB, Klaus. *A quarta revolução industrial*. Tradução de Daniel Moreira Miranda. São Paulo: Edipro, 2016, p. 115. O autor comenta: “Na quarta revolução industrial, a conectividade digital possibilitada por tecnologias de *software* está mudando profundamente a sociedade. A escala do impacto e a velocidade das mudanças fazem que a transformação seja diferente de qualquer outra revolução industrial da história da

*sociedade em rede*²⁰ – já conectada – em uma sociedade hiperconectada, potencializada pela imediatez das comunicações²¹. A viabilização dessa transição depende de tecnologias como o 5G, em um contexto marcado pela virtualização escalável da Internet das Coisas (*Internet of Things*, ou *IoT*). Todavia, não se pode olvidar de que “a IoT pode ser vista em diferentes dimensões pelos diferentes setores da academia e da indústria; qualquer que seja o ponto de vista, a IoT ainda não atingiu a maturidade e é vulnerável a todos os tipos de ameaças e ataques”²².

Então, a resposta à primeira dúvida é negativa. Um sistema de inteligência artificial, ainda que supere o “Teste de Turing” ou qualquer outro similar, não é capaz de atingir a capacidade cognitiva humana²³.

Isso não significa, contudo, que se deva desconsiderar que a turboglobalização já modificou o modo de vida contemporâneo, ao passo que o Estado, incapaz de responder com celeridade às demandas e contingências desveladas pelas inovações incessantes, se vê desafiado a editar marcos regulatórios.

Os direitos humanos são os mais impactados, uma vez que, por meio de uma compreensão crítica, é possível identificar, para além dos direitos formalmente reconhecidos, a surgimento de novos direitos atrelados à satisfação de necessidades humanas como ‘fonte permanente de direitos insurgentes’²⁴, encontrando barreiras desde a concepção de Herrera Flores sobre os ‘processos de luta pela dignidade’²⁵ e combatendo o universalismo abstrato dos direitos humanos para reinventá-los a partir de uma racionalidade de resistência²⁶.

O futuro distópico – quiçá iminente – que se imagina ao tratar de novas tecnologias impõe novas compreensões sobre fatos sociais que ultrapassam a dogmática jurídica. E, embora não se possa dizer que sistemas de IA são inteligentes ou capazes de pensar como humanos, pois não o são, também não se pode negar que eventos danosos podem decorrer de falhas desses

humanidade. O Conselho da Agenda Global do Fórum Econômico Mundial sobre o futuro do *Software* e da Sociedade realizou uma pesquisa com 800 executivos para avaliar quando os líderes empresariais acreditariam que essas tecnologias revolucionárias poderiam chegar ao domínio público em grau significativo e para compreender plenamente as implicações dessas mudanças para indivíduos, organizações, governo e sociedade”.

²⁰ A expressão ‘sociedade em rede’ foi originalmente proposta por van Dijk, em obra publicada originalmente em 1991. VAN DIJK, Jan. *The network society*. 2. ed. Londres: Sage, 2006, p. 6. Posteriormente, o conceito foi aprofundado pelos estudos de Castells, inclusive sob a ótica do direito. Para maiores detalhes, ver CASTELLS, Manuel. *The rise of the network society*. The information age: economy, society, and culture 2. ed. Oxford/West Sussex: Wiley-Blackwell, 2010, v. 1, p. 500.

²¹ Cf. MUNTADAS, Borja. *Imediatez*. Capitalismo y vidas aceleradas. Barcelona: Chiado Editorial, 2016.

²² PEPPE, Scott R. Regulating the Internet of Things: first steps toward managing discrimination, privacy, security, and consent. *Texas Law Review*, Austin, v. 93, p. 85-176, 2014, p. 147 *et seq.*

²³ Nesse sentido, conferir, por todos, CÍCUREL, Ronald; NICOLELIS, Miguel. *O cérebro relativístico: como ele funciona e por que ele não pode ser simulado por uma máquina de Turing*. Natal/Montreux/Durham/São Paulo: Kios Press, 2015.

²⁴ WOLKMER, Antonio Carlos. As Necessidades Humanas como Fonte Permanente de Direitos Insurgentes. *Veredas do Direito: Direito Ambiental e Desenvolvimento Sustentável*, Belo Horizonte, v. 1, n. 3, p.85-92, 2004

²⁵ HERRERA FLORES, Joaquín. Entrevista com Joaquim Herrera Flores. *Constituição e Democracia*, Brasília, UnB, jun. 2008, p. 12-13.

²⁶ HERRERA FLORES, Joaquín. *A (re)invenção dos direitos humanos*. Tradução de Carlos Roberto Diogo Garcia, de Antônio Henrique Graciano Suxberger e de Jefferson Aparecido Dias. Florianópolis: Fundação Boiteux, 2009, p. 21, 152 e 160.

sistemas, se convolvendo em ilícitos merecedores de tutela jurídica.

3. ALGUMAS SINALIZAÇÕES REGULATÓRIAS: RESPONSABILIDADE CIVIL E DANOS POR FALHAS DE ALGORITMOS

Embora sofisticados, os algoritmos que fazem funcionar sistemas de inteligência artificial, inclusive os que são baseados em técnicas de aprendizado profundo (*deep learning*), ainda operam no nível da predição heurística, ou seja, são dependentes de fórmulas matemáticas que indicam possibilidades para instruir decisões em função de análise probabilística²⁷. Por esse motivo, não são inteligentes na acepção biopsicológica que tal adjetivo carrega, e o acalorado debate em torno da atribuição de “personalidade eletrônica” a máquinas – aventado na União Europeia há alguns anos – não se apresenta como solução única, muito menos definitiva, ao problema, embora seja importante revisitá-lo. Bem ao contrário, é tema muito criticado e mal encarado pela doutrina, ainda que seja cogitado apenas a nível especulativo.

Como anota Wolfgang Hoffmann-Riem, “ao se lidar com a IA no direito, é preciso superar reduções até agora costumeiras no modo de abordar a questão, como, por exemplo, o foco nos potenciais de vigilância”²⁸. Aliás, a própria nomenclatura “inteligência artificial”, que foi cunhada por John McCarthy com a nobre intenção de simplificar futuras pesquisas sobre o tema pela aglutinação de resultados em um único verbete²⁹, embora tenha se popularizado como um gênero³⁰, não reflete o atual estado da técnica, pois a natureza matemática das predições conduz à conclusão de Vernor Vinge quanto ao não atingimento da ‘singularidade tecnológica’³¹, isto é, do momento em que o avanço tecnológico propiciará verdadeira simbiose – e possível indistinção – entre o biológico e o tecnológico, permitindo a superação do ‘jogo da imitação’ de

²⁷ KELLEHER, John D.; MAC NAMEE, Brian; D'ARCY, Aiofe. *Fundamentals of machine learning for predictive data analytics: algorithms, worked examples, and case studies*. Cambridge: The MIT Press, 2015, p. 1-16.

²⁸ HOFFMANN-RIEM, Wolfgang. Inteligência artificial como oportunidade para a regulação jurídica. Tradução de Luiz Sander; revisão de Laura Schertel Mendes. *Revista Direito Público*, Porto Alegre, v. 16, n. 90, p. 11-38, nov./dez. 2019, p. 13.

²⁹ MCCARTHY, John. *Artificial Intelligence (complete)* – Thinking Allowed [Entrevistador: Jeffrey Mishlove]. YouTube (video). 3 nov. 2011. Disponível em: <https://www.youtube.com/watch?v=Ozipf13jRr4>. Acesso em: 18 jun. 2023.

³⁰ Na cartilha de princípios sobre inteligência artificial da OCDE, por exemplo, é explicitada a preferência terminológica por “sistemas de inteligência artificial”, deixando claro que não se está a definir postulados para o complexo tema da ‘inteligência’ ou sobre os limites do pensar e do agir intuitivo inerentes ao ser humano quando se analisa a propensão do desenvolvimento algorítmico: “An AI system is a machine-based system that is capable of influencing the environment by producing an output (predictions, recommendations or decisions) for a given set of objectives. It uses machine and/or human-based data and inputs to (i) perceive real and/or virtual environments; (ii) abstract these perceptions into models through analysis in an automated manner (e.g., with machine learning), or manually; and (iii) use model inference to formulate options for outcomes. AI systems are designed to operate with varying levels of autonomy”. ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. *Artificial intelligence Principles*. Recommendation on the Council on Artificial Intelligence. 2019. Disponível em: <https://oecd.ai/en/ai-principles> Acesso em: 18 mar. 2023.

³¹ VINGE, Vernor. The coming technological singularity: How to survive in the post-human era. In: *Interdisciplinary Science and Engineering in the Era of Cyberspace*. NASA John H. Glenn Research Center at Lewis Field, Cleveland, 1993, p. 11-22. Disponível em: <https://ntrs.nasa.gov/search.jsp?R=19940022856> Acesso em: 18 mar. 2023.

Turing³².

Insufismavelmente, a indagação sobre se um sistema de IA é capaz “pensar” e, em última análise, causar danos pela prática de “discriminações algorítmicas” possui resposta veementemente negativa!

Não obstante, certa polêmica surgiu pelo próprio caráter prospectivo e especulativo do debate em torno da famigerada “personalidade eletrônica”, mas o tema, em si, não é novo. Em 1992, Lawrence Solum publicou artigo pioneiro sobre essa questão³³, muito antes do avanço vertiginoso da tecnologia no século XXI. Fato é que as inquietações ganharam forças a partir de 2015, quando a União Europeia publicou o *Draft Report with recommendations on civil law rules and robotics* (2015/2103), no qual era possível notar algumas preocupações quanto aos danos causados por sistemas algorítmicos automatizados.

A partir desse documento, em 16 de fevereiro de 2017, o Parlamento Europeu aprovou uma Resolução (“Disposições de Direito Civil sobre Robótica”) que, expressamente e em caráter prospectivo, prevê em sua diretriz 59, “f”, a possibilidade de atribuição de personalidade eletrônica a robôs.³⁴ Segundo Mafalda Miranda Barbosa, a proposta revela a necessidade de estruturação de um futuro regulamento em torno dos princípios da precaução, da reversibilidade, da segurança e da responsabilidade.³⁵ Isso porque, em linha com o que descrevem Marcos Ehrhardt Jr. e Gabriela Buarque Silva, indubitavelmente, “caracterizar a personalidade como um atributo jurídico não implica dizer, contudo, que o legislador possui ampla liberdade para instituir ou destituir personalidades, especialmente no que tange ao ser humano”³⁶.

O conceito de singularidade tecnológica também aparece nas pesquisas de Ray Kurzweil³⁷ e indica um momento de emancipação. Em simples termos, os algoritmos se

³² TURING, Alan M. On computable numbers, with an application to the *Entscheidungsproblem*. *Proceedings of the London Mathematical Society*, Londres, v. 42, n. 1, p. 230-265, nov. 1936. A ideia de uma máquina capaz de pensar leva a reflexões sobre os clássicos estudos de Alan Turing acerca do *Entscheidungsproblem* (dilema da tomada de decisão). Porém, o que Turing desconsiderou neste estudo foi justamente a possibilidade do ardid, da mentira e da dissimulação – comportamentos tipicamente humanos – e, ao delinear as premissas de seu famoso “Teste de Turing”, notou que potenciais erros de identificação poderiam surgir. Seria necessário avançar rumo à consolidação de um modelo mais complexo, por ele batizado de “Jogo da Imitação”, para que fosse possível aferir o potencial de uma máquina que pretenda, de forma convincente, “se passar por humana”.

³³ SOLUM, Lawrence. Legal personhood for Artificial Intelligences. *North Carolina Law Review*, Chapel Hill, v. 70, n. 4, p. 1231-1287, 1992.

³⁴ Com efeito: “59. Insta a Comissão a explorar, analisar e ponderar, na avaliação de impacto que fizer do seu futuro instrumento legislativo, as implicações de todas as soluções jurídicas possíveis, tais como: (...) f) Criar um estatuto jurídico específico para os robôs a longo prazo, de modo a que, pelo menos, os robôs autônomos mais sofisticados possam ser determinados como detentores do estatuto de pessoas eletrônicas responsáveis por sanar quaisquer danos que possam causar e, eventualmente, aplicar a personalidade eletrônica a casos em que os robôs tomam decisões autônomas ou em que interagem por qualquer outro modo com terceiros de forma independente”. PARLAMENTO EUROPEU. *Resolução de 16 de fevereiro de 2017*. Disposições de Direito Civil sobre Robótica. Disponível em: https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_PT.html. Acesso em: 18 mar. 2023.

³⁵ BARBOSA, Mafalda Miranda. Inteligência artificial, e-persons e direito: desafios e perspectivas. *Revista Jurídica Luso-Brasileira*, Lisboa, ano 3, n. 6, p. 1475-1503, 2017, p. 1501-1502.

³⁶ EHRHARDT JÚNIOR, Marcos; SILVA, Gabriela Buarque Pereira. Pessoa e sujeito de direito: reflexões sobre a proposta europeia de personalidade jurídica eletrônica. *Revista Brasileira de Direito Civil*, Belo Horizonte, v. 23, n. 1, p. 57-79, jan./mar. 2020, p. 61.

³⁷ KURZWEIL, Ray. *The age of spiritual machines: when computers exceed human intelligence*. Nova York: Viking, 1999. p. 213.

tornariam tão complexos, em razão do avanço do *machine* e do *deep learning*, que ultrapassariam os limites da predição estatística e do singelo suporte à tomada de decisões em função da heurística e, efetivamente, se tornariam máquinas capazes de ‘pensar’.

Jean-Gabriel Ganascia explica a proposição de Kurzweil:

De acordo com Kurzweil, muito em breve estaremos baixando nossa consciência para máquinas, o que nos fornecerá uma forma de imortalidade. Isso resultará inevitavelmente da aceleração do progresso. Com efeito, segundo ele, a lei geral a que obedece toda forma de evolução, seja o desenvolvimento biológico, o aperfeiçoamento das civilizações ou o aperfeiçoamento das tecnologias, é de natureza intrinsecamente exponencial.³⁸

Essa mudança de paradigma permitiria aos algoritmos – que se tornariam, a partir de então, ‘singulares’, ‘artificialmente inteligentes’ ou dotados de ‘inteligência artificial geral’, e propensos à tomada de decisões morais – agir como um humano, com pensamento abstrato, discernimento baseado em valores e potencial discriminatório decorrente da formação de uma personalidade moldada e robustecida pelo acúmulo informacional³⁹.

Por isso, a adoção de nomenclaturas específicas tem seu valor, como anota Visa Kurki:

Certas formas emergentes de personalidade jurídica provavelmente exigirão novas nomenclaturas. Se a personalidade jurídica for estendida à IA ou a animais, surgirá a questão de ‘como chamar’ essas formas de personalidade jurídica. Novos termos ou frases, como ‘personalidade animal’ ou ‘personalidade eletrônica’, podem ser facilmente introduzidos. Tais soluções poderiam resolver possíveis questões de interpretação jurídica do conceito, da mesma forma que ‘pessoa natural’ e ‘pessoa jurídica’ são usadas, hoje, quando diferentes escopos de personalização precisam ser mantidos distintos.⁴⁰

Mafalda Miranda Barbosa ressalta que sistemas de IA não podem ser considerados *e-persons*, pois ostentam autonomia “puramente ‘tecnológica, fundada nas potencialidades da combinação algorítmica que é fornecida ao *software*. Está, portanto, longe do agir ético dos humanos, em que radica o ser pessoa”⁴¹. Ainda tratando do tema, Marcelo de Oliveira Milagres afirma que “parece que o fundamento para o reconhecimento da personalidade às máquinas

³⁸ GANASCIA, Jean-Gabriel. *Le mythe de la singularité: faut-il craindre l'intelligence artificielle?* Paris: Éditions du Seuil, 2017, p. 13, tradução livre. No original: “Selon Kurzweil, nous téléchargerons très bientôt notre conscience sur des machines, ce qui nous procurera une forme d'immortalité. Cela résultera fatalement de l'accélération des progrès. En effet, d'après lui, la loi générale à laquelle obéit toute forme d'évolution, qu'il s'agisse du développement biologique, du perfectionnement des civilisations ou de l'amélioration des technologies, est de nature intrinsèquement exponentielle.”

³⁹ MADSBJERG, Christian. *Sensemaking: the power of the humanities in the age of the algorithm*. Nova York: Hachette, 2017, *E-book*, Cap. VI.

⁴⁰ KURKI, Visa A. J. *A theory of legal personhood*. Oxford: Oxford University Press, 2019, p. 200, tradução livre. No original: “Certain emerging forms of legal personhood will likely require new nomenclature. If legal personhood is extended to AIs or animals, the question will arise what to call these forms of legal personhood. New terms or phrases, such as ‘animal person’ or ‘electronic person’, could easily be introduced. Such solutions could resolve potential questions of statutory interpretation, much as ‘natural person’ and ‘artificial person’ are used today when different classes of legal persons need to be kept distinct”.

⁴¹ BARBOSA, Mafalda Miranda. Inteligência artificial, e-persons e direito: desafios e perspectivas. *Revista Jurídica Luso-Brasileira*, Lisboa, ano 3, n. 6, 2017, p. 1482.

autônomas é a maior viabilidade do ressarcimento dos danos decorrentes de suas ações⁴², a indicar a prevalência de um argumento pragmático e não necessariamente adaptado às peculiaridades do desenvolvimento tecnológico.

De fato, é indubitável que a singularidade tecnológica não foi atingida, mas propostas que buscam investigar sua compatibilidade com institutos jurídicos acabam refutando a ideia de atribuição da “personalidade eletrônica” por completo. A primeira alternativa apontada pela doutrina, de cariz intermediário, surgiu na doutrina alemã, em estudo de Jan-Erik Schirmer⁴³ sobre a *Teilrechtsfähigkeit*, publicado em 2020. A sugestão do autor é considerar um robô parcialmente capaz, atribuindo-lhe “personalidade jurídica fragmentária” (quase indistinta da capacidade jurídica) construída evolutivamente⁴⁴, ou seja, estabelecendo-se níveis de maior ou menor capacidade jurídica até o atingimento da singularidade tecnológica.

Em síntese, até que a mencionada singularidade tecnológica seja atingida, modelos pouco usuais podem servir para tutelar provisoriamente situações jurídicas – especialmente as de cariz existencial –, evitando dubiedades.

Para ações discriminatórias, assim como na preocupação que se tem com os vieses algorítmicos, a inquietação é ainda mais acentuada, pois se poderia exigir de um ‘algoritmo singular’ maior ou menor previsibilidade quanto aos danos que pudesse causar, e graus de risco escalonados poderiam ser estabelecidos e impostos aos desenvolvedores. Nos dizeres de Caitlin Mulholland, “para que uma pessoa seja obrigada a reparar um dano injusto, é fundamental que ela tenha a autonomia de atuação, isto é, tenha a capacidade de reconhecer a licitude ou ilicitude de sua conduta e, ao mesmo tempo, a habitualidade de identificar e prever a potencialidade danosa desta”⁴⁵.

Por esse motivo, “parece ser excessivamente artificial – ou mera ficção científica – imaginar um robô androide perfeitamente inteligente, andando pelas ruas sem qualquer controle”⁴⁶. Não se trata disso, mas há inúmeros exemplos já colhidos da experiência recente em

⁴² MILAGRES, Marcelo de Oliveira. A robótica e as discussões sobre a personalidade eletrônica. In: EHRHARDT JÚNIOR, Marcos; CATALAN, Marcos; MALHEIROS, Pablo (coord.). *Direito civil e tecnologia*. Belo Horizonte: Fórum, 2020, p. 513.

⁴³ SCHIRMER, Jan-Erik. Artificial Intelligence and legal personality. “Teilrechtsfähigkeit”: A partial legal status made in Germany. In: WISCHMEYER, Thomas; RADEMACHER, Timo (ed.). *Regulating Artificial Intelligence*. Cham: Springer, 2020, p. 134.

⁴⁴ FALEIROS JÚNIOR, José Luiz de Moura; MENKE, Fabiano. “Teilrechtsfähigkeit”: uma proposta alemã para a responsabilização civil na IA. *Migalhas de Responsabilidade Civil*, 06 ago. 2020. Disponível em: <https://s.migalhas.com.br/S/8AF9D>. Acesso em: 18 mar. 2023. Com efeito: “(...) Schirmer descreve a personalidade jurídica ostentada por humanos como um ‘pote de doces’ que está cheio desde o começo; o pote representaria a personalidade jurídica e os doces simbolizariam direitos específicos, logo, um pote cheio de doces indicaria a personalidade em sua plenitude: ou se tem o pote cheio, ou não há pote algum. Na *Teilrechtsfähigkeit*, a diferença adviria da atribuição de personalidade, mas sem direitos pré-concebidos pelo ordenamento; o pote existiria, mas estaria inicialmente vazio de doces, sendo preenchido, pouco a pouco, em sintonia com a própria evolução da personalidade, até que se tornasse plena”.

⁴⁵ MULHOLLAND, Caitlin. Responsabilidade civil e processos decisórios autônomos em sistemas de Inteligência Artificial (IA): autonomia, imputabilidade e responsabilidade. In: FRAZÃO, Ana; MULHOLLAND, Caitlin (coord.). *Inteligência Artificial e Direito: ética, regulação e responsabilidade*. São Paulo: Thomson Reuters Brasil, 2019, p. 332.

⁴⁶ TOMASECIVILUS FILHO, Eduardo. Inteligência Artificial e direitos da personalidade: uma contradição em termos? *Revista da Faculdade de Direito da Universidade de São Paulo*, São Paulo, v. 113, p. 133-149, jan./dez. 2018, p. 142.

que os vieses – total ou parcialmente previsíveis – acarretaram algum tipo de dano, inclusive de natureza enviesada, conforme o seguinte exemplo relatado:

Exemplo curioso é o da tecnologia *LiDAR* (acrônimo de *Light Radar*), baseada no rastreamento da luz refletida por objetos do entorno de um veículo autônomo. A luz que retorna é absorvida por sensores para o cálculo de sua densidade (luminância), que, catalogada, alimenta de dados um algoritmo capaz de processar e projetar, em plano tridimensional, obstáculos presentes nas cercanias, deles desviando para prevenir abalroamentos. Dessa forma, utilizando-se de feixes de luz, milhões de cálculos são processados em segundos para permitir o deslocamento desses veículos sem qualquer intervenção humana, e, a partir do aprendizado de máquina (*machine learning*), vão se tornando mais confiáveis.⁴⁷

A mencionada tecnologia vem sendo empregada para o desenvolvimento de carros, navios e *drones* autônomos. Porém, notícias recentes dão conta da maior propensão desses veículos ao atropelamento de pessoas negras⁴⁸ pelo fato de o feixe de luz provocar enviesamento, na medida em que pessoas negras deixam de ser identificadas nas cercanias pelo simples fato de toda a tecnologia ser dependente da luminância aferida pelo reflexo, e é sabido que cores mais escuras absorvem mais luz. Eis o viés, totalmente previsível e representativo de riscos que podem ser evitados.

Embora já seja possível afirmar uma ‘branquidão’ das tecnologias baseadas em inteligência artificial⁴⁹, um exemplo como o do sensor *LiDAR* não impõe dizer que o algoritmo é racista, pois há, em verdade, um viés a ser considerado na leitura que se faz do resultado concreto. O problema é, pois, mais complexo do que parece e demanda reflexões profundas para a compatibilização dessa e de outras tecnologias potencialmente propensas à causação de danos discriminatórios em momento no qual não se atingiu a singularidade.

Surge, então, uma indagação complementar: como diferenciar o preconceito dos vieses? Por certo, se a singularidade tecnológica ainda não foi atingida e não se tem sistemas efetivamente inteligentes, não se pode afirmar que algoritmos pensam, tampouco que acumulam vivências, impressões ou percepções ao longo das etapas de treinamento (no curso do *machine learning*) a que são submetidos.

Se limitam, como foi dito, a processar dados e partir da heurística. Humanos, por outro lado, estão sujeitos ao enviesamento de suas percepções, mas, em razão de sua inerente

⁴⁷ FALEIROS JÚNIOR, José Luiz de Moura. Discriminação por algoritmos de Inteligência Artificial: a responsabilidade civil, os vieses e o exemplo das tecnologias baseadas em luminância. *Revista de Direito da Responsabilidade*, Coimbra, ano 2, p. 1007-1043, 2020, p. 1009.

⁴⁸ CUTHBERTSON, Anthony. Self-driving cars more likely to drive into black people, study claims. *The Independent*, 06 mar. 2019. Disponível em: <https://www.independent.co.uk/life-style/gadgets-and-tech/news/self-driving-car-crash-racial-bias-black-people-study-a8810031.html>. Acesso em: 18 mar. 2023; SAMUEL, Sigal. A new study finds a potential risk with self-driving cars: failure to detect dark-skinned pedestrians. *Vox*, 06 mar. 2019. Disponível em: <https://www.vox.com/future-perfect/2019/3/5/18251924/self-driving-car-racial-bias-study-autonomous-vehicle-dark-skin>. Acesso em: 18 mar. 2023.

⁴⁹ CAVE, Stephen; DIHAL, Kanta. The whiteness of AI. *Philosophy & Technology*, Londres, v. 33, p. 685-703, 2020, p. 698. Destacam: “While we believe the attribution to AI of these qualities, so strongly associated with Whiteness, goes a long way to making sense of the racialisation of anthropomorphic intelligent machines, we also want to propose one further hypothesis: that the Whiteness of the machines allows the White utopian imagination to fully exclude people of colour”.

racionalidade, podem construir visões e opiniões verdadeiramente preconceituosas⁵⁰.

A nível estrutural, todas as distorções preconceituosas existentes em sociedade podem ser projetadas para algoritmos funcionalizados a partir de conjuntos de dados que se tornam imprescindíveis para o trabalho heurístico. Noutros termos, os mesmos vieses que existem no pensamento humano podem se traduzir em vieses algorítmicos e em nefasta discriminação⁵¹. Não se trata de dizer, porém, que o algoritmo discrimina; se trata de dizer que suas análises preditivas são contaminadas por conjuntos de dados previamente enviesados.

O preconceito se difere dos vieses a partir do que Daniel Kahneman e Amos Tversky descrevem como “psicologia da predição”. Segundo os autores, há duas estruturas preditivas essenciais para que se realize qualquer tentativa de tomada de decisão: (i) a *predição por categorias*, em que a previsão é dada na forma nominal, por exemplo, o vencedor em uma eleição, o diagnóstico de um paciente pela avaliação clínica ou a ocupação futura de uma pessoa a partir de testes a partir de testes de aptidão; (ii) a *predição por números*, e, que a previsão é dada na forma numérica, por exemplo, o valor futuro de uma determinada ação ou a média de notas de um aluno⁵². Algoritmos funcionam a partir da segunda estrutura preditiva, pois são dependentes dos números para realizar inferências matemáticas baseadas em análise probabilística⁵³. Suas “decisões” são norteadas por dados prévios, mesmo nas técnicas mais avançadas, que conseguem valorar tais dados em estruturas multicamadas, antes de processá-los – atribuindo-lhes um “peso”, como se faz em algoritmos de retropropagação de erro (típicos do *deep learning*)⁵⁴ –, e a escolha final ainda é baseada em números que formam uma rede neural artificial. Não há, portanto, influência acrítica das percepções do mundo ou das vivências.

4. VIÉS RACIAL E RESPONSABILIDADE CIVIL A PARTIR DE FALHAS ALGORÍTMICAS: POSSÍVEIS SOLUÇÕES

A compreensão do racismo estrutural no século XXI incita necessariamente uma análise do âmbito social, mormente das relações históricas, políticas e econômicas, bem como do

⁵⁰ Cf. EBERHARDT, Jennifer. *Biased: uncovering the hidden prejudice that shapes what we see, think, and do*. Nova York: Viking, 2019.

⁵¹ NOBLE, Safiya Umoja. *Algorithms of oppression: how search engines reinforce racism*. Nova York: NYU Press, 2018, p. 141. Comenta: “Knowledge management reflects the same social biases that exist in society, because human beings are at the epicenter of information curation. These practices of the past are part of the present, and only committed and protracted investments in repairing knowledge stores to reflect and recenter all communities can cause a shift toward equality and inclusion in the future. This includes reconciling our brutal past rather than obscuring or minimizing it. In this way, we have yet to fully confront our histories and reconstitute libraries and museums toward reconciliation and reparation”.

⁵² KAHNEMAN, Daniel; TVERSKY, Amos. On the psychology of prediction. In: KAHNEMAN, Daniel; SLOVIC, Paul; TVERSKY, Amos (ed.). *Judgement under uncertainty: heuristics and biases*. Cambridge: Cambridge University Press, 1982, p. 48.

⁵³ SUNSTEIN, Cass R. *Valuing life: humanizing the regulatory state*. Chicago: The University of Chicago Press, 2014. p. 137. O autor comenta: “When people use simple heuristics, or mental shortcuts, it is generally because they work well, in the sense that they enable us to make good decisions. But even if heuristics usually work well, they can lead to big errors. When we make inaccurate assessments of probabilities, it may well be because simple heuristics are leading us astray”.

⁵⁴ KELLEHER, John D. *Deep learning*. Cambridge: The MIT Press, 2019, p. 223.

entendimento dos conceitos de raça e racismo⁵⁵. A noção de raça, em sua origem, está interligada à ação humana de criar agrupamentos de espécies biológicas semelhantes entre si⁵⁶; a transposição do termo para uso em agrupamentos humanos surgiu no século XVI⁵⁷ e se popularizou no século XVIII⁵⁸. O termo raça constitui sentido a partir do *locus* de emanção em que está inserido ou sob o qual é analisado, sendo, portanto, um conceito dinâmico⁵⁹ que deve ser analisado a partir de seu contexto⁶⁰.

Considerando o cenário do século XVI, o conceito de raça ganhou seu conteúdo por meio do processo de expansão marítima europeia e a descoberta da existência de novos povos caracteristicamente distintos dos europeus⁶¹. É com o surgimento do iluminismo e a consequente centralização do homem, que se cria a noção distintiva entre o ser civilizado – baseado na concepção de homem branco europeu – e o selvagem⁶². Em nome da imposição civilizatória aos não-ocidentais, inúmeras batalhas foram travadas num processo violento de imposição da razão universal denominado colonialismo⁶³. Nesse contexto, a classificação racial dos indivíduos fora utilizada como ferramenta de inferiorização dos povos colonizados⁶⁴, considerados a quintessência do mal: ausentes de ética e valores⁶⁵. O homem colono, na imposição de seus dogmas absolutos, retira o caráter humano da cultura dos povos colonizados para lhes estigmatizar como bárbaros e desalmados⁶⁶.

A noção de raça, no contexto europeu, estava frequentemente associada à categorização de grupos sociais com base, primordialmente, na cor da pele, ademais de existirem variadas outras diferenças físicas entre os seres humanos⁶⁷; portanto, a ideia de raça é socialmente construída em diferentes contextos históricos a partir da seleção de determinadas características

⁵⁵ ALMEIDA, Silvio Luiz de. *Racismo estrutural*. São Paulo: Pólen, 2019, p. 16.

⁵⁶ A palavra raça deriva do termo latim *ratio*, que remete à noção de ordem cronológica, podendo ser compreendida como o complexo de traços característicos – externos e internos – compartilhados por gerações. MENDES, Maria Manuela. Raça e racismo: controvérsias e ambiguidades / Race and racism: controversies and ambiguities. *Vivência: Revista de Antropologia*, v. 1, n. 39, 8 maio 2012, p. 102.

⁵⁷ ALMEIDA, Silvio Luiz de. *Racismo estrutural*. São Paulo: Pólen, 2019, p. 16.

⁵⁸ Durante o século XVIII, a teoria denominada “determinismo climático” ganhou popularidade procurando explicar as diferenças humanas como resultado fático de aspectos de localização geográfica e variáveis climáticas, fatores capazes de retardar o avanço civilizatório através da criação de variáveis físicas. MENDES, Maria Manuela. Raça e racismo: controvérsias e ambiguidades / Race and racism: controversies and ambiguities. *Vivência: Revista de Antropologia*, v. 1, n. 39, 8 maio 2012, p. 102.

⁵⁹ O conceito de raça é abordado por Almeida como não estático, adquirindo conteúdo a partir das circunstanciais sociais de determinado momento histórico. ALMEIDA, Silvio Luiz de. *Racismo estrutural*. São Paulo: Pólen, 2019, p. 18.

⁶⁰ BORGES, Gustavo Silveira; LIMA, Fernanda da Silva. Publicidade e racismo reverso: o que uma campanha publicitária tem a revelar sobre o racismo no Brasil. *Revista de Direito do Consumidor*, São Paulo, v. 123, p. 37-76, 2019.

⁶¹ ALMEIDA, Silvio Luiz de. *Racismo estrutural*. São Paulo: Pólen, 2019, p. 18.

⁶² ALMEIDA, Silvio Luiz de. *Racismo estrutural*. São Paulo: Pólen, 2019, p. 19.

⁶³ MBEMBE, Achille. *Crítica da razão negra*. São Paulo: N-1, 2018, p. 170.

⁶⁴ A razão universal fora imposta aos povos não ocidentais sob o discurso da Modernidade consistente num processo de autodescrição europeia como civilização desenvolvida. DUSSEL, Enrique. Europa, modernidade e eurocentrismo. In: LANDER, Edgardo (org.). *A colonialidade do saber: eurocentrismo e ciências sociais, perspectivas latinoamericanas*. São Paulo: CLACSO, 2005, p. 28-29.

⁶⁵ FANON, Frantz. *Os condenados da terra*. Tradução de Enilce Albergaria Rocha e Lucy Magalhães. Juiz de Fora: Editora da UFJF, 2005, p. 36.

⁶⁶ FANON, Frantz. *Os condenados da terra*. Tradução de Enilce Albergaria Rocha e Lucy Magalhães. Juiz de Fora: Editora da UFJF, 2005, p. 38.

⁶⁷ MILES, Robert; BROWN, Malcolm. *Racism*. Londres: Routledge, 2003, p. 89.

com o intuito de servir aos interesses dominantes⁶⁸.

Já o racismo, derivado da ideia de raça⁶⁹, é “uma forma sistemática de discriminação que tem a raça como fundamento”⁷⁰. O racismo constitui a manifestação/concretização das relações e concepções existentes no cerne da sociedade; sua materialização em instituições físicas apenas reflete o comportamento já existente nas estruturas sociais⁷¹. Dessa forma, fala-se, portanto, que o racismo é estrutural.

É expressão da manutenção continuada das relações históricas de poder e dominação que acarretaram a exclusão social dos povos não ocidentais, denominados de “Outros”⁷². A concepção do Outro não se restringiu apenas à história do colonialismo, existindo antes mesmo da expansão europeia⁷³ – o Outro mulçumano, o Outro judeu, o Outro mulher durante a Idade Média, o Outro bárbaro –, mas é com a expansão marítima e o comércio de escravos africanos que o discurso do Outro africano como raça inferior centra-se na cor da pele.

Ao fenômeno de inferiorização dos povos indígenas e escravos africanos nas Américas através da utilização da raça, Maldonado-Torres denomina de heterogeneidade colonial⁷⁴. Apesar do fim do colonialismo com a emancipação das colônias da América no século XIX, a colonialidade ainda perdura até os dias atuais⁷⁵⁻⁷⁶. Como resultado histórico, a América Latina contava com aproximadamente 130 milhões de afrodescendentes em sua população em 2017, sendo, de acordo com a Comissão Econômica para a América Latina e o Caribe (CEPAL), a raça, uma das condições de perpetuação das desigualdades sociais⁷⁷.

Já o Brasil possui 56,10% de sua população autodeclarada negra, conforme Pesquisa Nacional por Amostra de Domicílios (Pnad) do IBGE no ano de 2019; contando com 209,2

⁶⁸ MILES, Robert; BROWN, Malcolm. *Racism*. Londres: Routledge, 2003, p. 89.

⁶⁹ MILES, Robert; BROWN, Malcolm. *Racism*. Londres: Routledge, 2003, p. 88.

⁷⁰ ALMEIDA, Silvio Luiz de. *Racismo estrutural*. São Paulo: Pólen, 2019, p. 23.

⁷¹ ALMEIDA, Silvio Luiz de. *Racismo estrutural*. São Paulo: Pólen, 2019, p. 32.

⁷² MILES, Robert; BROWN, Malcolm. *Racism*. Londres: Routledge, 2003, p. 51.

⁷³ Os povos greco-romanos denominavam de Outros àqueles para além de seus limites territoriais, conhecidos como povos bárbaros; durante a Idade Média, o Outro fora representado naqueles que divergiam da cultura cristã, principalmente a imagem da mulher indomável; já durante a expansão do cristianismo europeu, os povos do Oriente praticantes do islamismo eram vistos de maneira intimidadora e portanto, considerados o Outro mulçumano, forjando relações de violência e guerras à cultura do islã – acarretando a ocupação de suas terras – e a sua representação como “inimigo da cristandade”. MILES, Robert; BROWN, Malcolm. *Racism*. Londres: Routledge, 2003, p. 19-30.

⁷⁴ MALDONADO-TORRES, Nelson. Sobre la colonialidad del ser: contribuciones al desarrollo de un concepto. *El giro decolonial. Reflexiones para una diversidad epistémica más allá del capitalismo global*, 2007, p. 132-133.

⁷⁵ OLIVEIRA, Luiz Fernandes de. *Educação e militância decolonial*. Rio de Janeiro: Editora Selo Novo, 2018, p. 46.

⁷⁶ CORRÊA, Bianca Kremer Nogueira. *Direito e tecnologia em perspectiva amefricana: autonomia, algoritmos e vieses raciais*. 2021. 299p. Tese (Doutorado em Direito) – Faculdade de Direito, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2021, p. 265-266. Anota: “Na tecnologia, o racismo algorítmico é reproduzido por ferramentas de busca, plataformas digitais, bancos de dados e tecnologias de vigilância, por meio de sistemas opacos e sem nenhum tipo de obrigatoriedade de explicação sobre seu funcionamento, utilizados indiscriminadamente por uma sociedade cada vez mais digital. (...) O fenômeno dos vieses algorítmicos e a reprodução do racismo por esses sistemas sobre corpos e experiências não brancos (...) precisam levar em consideração as hierarquias de humanidade que por aqui se desenvolveram – e desenvolvem – e a responsabilidade das nossas instituições e nossos sistema de justiça nesse processo”.

⁷⁷ COMISSIÓN ECONÓMICA PARA AMÉRICA LATINA Y EL CARIBE (CEPAL). *Panorama Social de América Latina*, 2016. Santiago, 2017.

milhões de habitantes, desses, 117,411 milhões assumem-se pretos ou pardos⁷⁸. Apesar disso, a quantidade ou representatividade negra não é condição suficiente nem mantém relação de causa e efeito com a ausência do racismo⁷⁹.

Nesse contexto, pertinente a indagação de Anupam Chander:

Dada a persistência da discriminação racial e de gênero generalizada no século XXI, não deveríamos esperar que algoritmos frequentemente programados por programadores racistas e sexistas nos manipulem para aceitar decisões racistas e sexistas? Os programadores são propensos a manipular algoritmos para exacerbar a discriminação existente na sociedade? Por meia dúzia de razões, acredito que a resposta seja não⁸⁰.

O racismo corporifica-se nas desigualdades estruturais – política, econômica, social, jurídica – que impõem ao negro lugar de marginalização na sociedade⁸¹. O reflexo do racismo é visto nas desigualdades existentes no mercado de trabalho, nas condições precárias de moradia, na distribuição do rendimento, nas taxas de homicídio, nos níveis de educação; cerca de 32,9% das pessoas negras sobrevivem com rendimento abaixo do nível de pobreza no Brasil – equivalente ao valor de US\$ 5,50 diários – enquanto 15,4% da população branca vive abaixo dessa linha⁸². Já os dados sobre violência no país demonstram que mais de 75% das vítimas de homicídio são jovens negros⁸³. Por isso, compreender a diferença conceitual entre colonialismo e colonialidade é essencial para o entendimento das estruturas racistas que ainda permanecem no cotidiano das ex-colônias⁸⁴.

Considerando o cenário disposto, nota-se a existência de estratégias de invisibilidade histórica e cultural que não se pode desconsiderar⁸⁵; portanto, faz-se necessário o

⁷⁸ INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Pesquisa Nacional por Amostra de Domicílios (Pnad) Contínua*. 2019. Disponível em: <https://sidra.ibge.gov.br/tabela/6403> Acesso em: 18 mar. 2023.

⁷⁹ ALMEIDA, Silvio Luiz de. *Racismo estrutural*. São Paulo: Pólen, 2019, p. 33.

⁸⁰ CHANDER, Anupam. The racist algorithm? *Michigan Law Review*, Ann Arbor, v. 115, n. 6, p. 1023-1045, 2017, p. 1028, tradução livre. No original: "Given the persistence of widespread racial and gender discrimination in the twenty-first century, should we not expect algorithms often programmed by racist and sexist programmers to manipulate us towards accepting racist and sexist decisions? Are programmers likely to manipulate algorithms to exacerbate existing discrimination in society? For a half-dozen reasons, I believe the answer is no".

⁸¹ ALMEIDA, Silvio Luiz de. *Racismo estrutural*. São Paulo: Pólen, 2019, p. 34.

⁸² INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Desigualdades Sociais por Cor ou Raça no Brasil*. Estudos e Pesquisas: Informação Demográfica e Socioeconômica, n. 41, [S.D]. Disponível em https://biblioteca.ibge.gov.br/visualizacao/livros/liv101681_informativo.pdf Acesso em: 18 mar. 2023.

⁸³ INSTITUTO DE PESQUISA ECONÔMICA APLICADA. Atlas da violência. Brasília/Rio de Janeiro/São Paulo: IPEA, *Fórum Brasileiro de Segurança Pública*, 2019. v. 11, 2020, p. 47.

⁸⁴ Adota-se para isso, a diferenciação realizada por Maldonado-Torres e Aníbal Quijano; o colonialismo é baseado nas relações de dominação entre Império e colônia, ele remonta à exploração colonial e antecede a colonialidade que, possui um sentido mais amplo, referindo-se às diversas relações e a noção de raça enraizada no subjetivo das populações decorrentes do colonialismo. MALDONADO-TORRES, Nelson. Sobre la colonialidad del ser: contribuciones AL desarrollo de un concepto. In: CASTRO-GOMEZ, Santiago; GROSGOUEL, Ramón (org.). *El giro decolonial: reflexiones para una diversidad epistémica más Allá del capitalismo global*. Bogotá: Universidad Javeriana-Instituto Pensar/Universidad Central-IESCO/Siglo del Hombre Editores, 2007, p. 131; QUIJANO, Aníbal. Colonialidade do poder e classificação social. *Epistemologias do Sul*. São Paulo: Cortez, 2010, p. 93.

⁸⁵ MELO, Miguel Ângelo Silva de; BANDEIRA, João Adolfo Ribeiro. Crítica aos estereótipos e ideias racistas no Brasil sob o prisma dos estudos Pós-coloniais. *Revista Direito e Práxis*, v. 7, n. 15, 2016, p. 218.

desenvolvimento de um pensamento crítico para que releituras de situações danosas, que envolvem o instituto da responsabilidade civil em razão de vieses discriminatórios, possam ser adequadamente realizadas. Isso porque, de fato, não parece ser crível que preconceitos individuais (de programadores e desenvolvedores) sejam o núcleo de enviesamentos algorítmicos, mas, sim, que vieses podem resultar de distorções existentes na própria sociedade. A despeito disso, aspectos éticos devem nortear o desenvolvimento de sistemas de IA, uma vez que, a nível contextual, representações algorítmicas enviesadas podem ser altamente danosas⁸⁶.

Assim, toma-se a liberdade de transcrever os seis argumentos de Anupam Chander sobre as razões pelas quais é negativa a resposta a tantas inquietações em relação à propensão discriminatória por viés racial de sistemas de IA:

Primeiro, uma vez que grande parte da discriminação social é subconsciente ou inconsciente, é menos provável que seja codificada em algoritmos automatizados do que nos tomadores de decisão humanos que os algoritmos substituem. Grande parte da pesquisa recente sobre viés racial tem se concentrado em expor sua existência sem focar se é consciente ou não. (...) Não apenas o programador deve instruir o computador com grande precisão, mas as práticas modernas de programação também exigem que o programador documente (ou anote) o que o programa está fazendo. Devido a um processo de programação que requer tanto a escrita de instruções explícitas quanto a documentação do que um código específico faz, a discriminação inconsciente ou subconsciente é menos provável de se manifestar na programação de computadores do que na tomada de decisões humanas. (...)

Segundo, mesmo para programadores ou empresas que pretendem discriminar, o próprio processo de codificação provavelmente fará com que os programadores evitem realmente codificar a discriminação. Mesmo na ausência de divulgação forçada por meio de litígios, há o perigo de que uma discriminação codificada seja revelada posteriormente por *hackers* ou por *insiders* indignados com a discriminação. (...)

Terceiro, mesmo que as empresas se envolvam em outros tipos de conduta ilícita, isso não significa que elas provavelmente manipularão intencionalmente algoritmos para discriminar injustamente. (...)

Quarto, o processo deve ser comparado com a tomada não algorítmica de decisões, ou seja, a tomada de decisões por seres humanos. A caixa-preta (*black box*) final é a mente humana. Mesmo quando as decisões são tomadas por comitês, as deliberações geralmente não são registradas, exceto em atas altamente seletivas, e os membros do comitê muitas vezes se comprometem a manter as discussões em segredo. Preconceitos aplicados nesta caixa-preta nunca precisam ser registrados por escrito. (...)

Quinto, se os seres humanos agem com base em estereótipos formados por meio de um processo de discriminação estatística, algoritmos automatizados atuando em um ambiente de informações mais rico podem não estar sujeitos a uma discriminação estatística individualmente errônea semelhante. (...)

Por fim, preocupações com a caixa-preta da tomada de decisões humanas levaram a mudanças algorítmicas no passado. Quando o Facebook foi acusado de favorecer visões políticas liberais nas histórias de notícias identificadas como "tendências", a empresa buscou demonstrar claramente seu compromisso com a imparcialidade demitindo seus editores de notícias e substituindo-os completamente por um algoritmo automatizado⁸⁷.

⁸⁶ NOBLE, Safiya Umoja. *Algorithms of oppression: how search engines reinforce racism*. Nova York: NYU Press, 2018, p. 171.

⁸⁷ CHANDER, Anupam. The racist algorithm? *Michigan Law Review*, Ann Arbor, v. 115, n. 6, p. 1023-1045, 2017, p. 1028-1031, tradução livre. No original: "First, because much of societal discrimination is subconscious or unconscious, it is less likely to be encoded into automated algorithms than the human decisionmakers that the algorithms replace. Much of recent research into racial bias has moved toward

Segundo Laura Schertel Mendes e Marcela Mattiuzzo, “descrevemos algo como preconceituoso quando uma afirmação se baseia em generalizações estatísticas inconsistentes, mas também quando nos referimos a generalizações estatisticamente consistentes, mas não universais”⁸⁸. Para que não haja regresso, especialmente quanto à proteção jurídica que se deve conferir aos direitos humanos⁸⁹, caminhos devem ser mapeados para conciliar inovação e regulação. Sem dúvidas, o pluralismo jurídico global deve atuar como vetor da função promocional dos direitos humanos⁹⁰, abrindo espaços à tutela subjacente-valorativa da pessoa, mesmo em um ambiente permeado pela disrupção tecnológica.

O atingimento desse desiderato, na esteira do que defende Gustavo Zagrebelsky, somente ocorrerá se determinadas condicionantes estruturais se fizerem presentes, das quais a “*ductibilidade*” (maleabilidade) dos ordenamentos jurídicos constitucionais é a mais relevante, pois propicia o pacifismo e a integração democrática “através da rede de valores e procedimentos comunicativos que é, ademais, a única visão possível e não catastrófica da política em nosso tempo”⁹¹.

exposing its existence without focusing on whether it is conscious or not. (...) Not only must the programmer instruct the computer with great precision, but modern programming practices also require the programmer to document (or annotate) what the program is doing. Because of a programming process that requires both writing down explicit instructions and documenting what particular code does, unconscious or subconscious discrimination is less likely to manifest itself in computer programming than in human decisionmaking. (...) Second, even for programmers or companies who intend to discriminate, the process of coding itself is likely to cause programmers to shy away from actually encoding the discrimination. Even absent compelled disclosure through litigation, there is the danger that a hard-coded discrimination will be revealed later by hackers or by insiders disgusted by the discrimination. (...) Third, even if corporations engage in other kinds of wrongdoing, that does not mean that they are likely to intentionally manipulate algorithms to invidiously discriminate. (...) Fourth, the process must be compared with non-algorithmic decisionmaking—that is, decisionmaking by human beings. The ultimate black box is the human mind. Even where decisions are made by committee, the deliberations are typically not recorded except in highly selective minutes, and the members of the committee often pledge to hold discussions secret. Prejudices acted upon in this black box never have to be written down. (...) Fifth, if human beings act on stereotypes formed through a process of statistical discrimination, automated algorithms acting on a richer information environment may not be subject to similar individually erroneous statistical discrimination. (...) Finally, worries about the black box of human decisionmaking have led to algorithmic turns in the past. When Facebook was accused of favoring liberal political views in the news stories it identified as “trending,” it sought to clearly demonstrate its commitment to fairness by firing its news editors and replacing them fully with an automated algorithm”.

⁸⁸ MENDES, Laura Schertel; MATTIUZZO, Marcela. Discriminação algorítmica: conceito, fundamento legal e tipologia. *Revista Direito Público*, Porto Alegre, v. 16, n. 90, p. 39-64, nov./dez. 2019, p. 49.

⁸⁹ José de Oliveira Ascensão aduz que os direitos do homem (direitos humanos, em sentido amplo), quando positivados em documentos internacionais de proteção e promoção da pessoa humana são considerados direitos humanos; quando positivados nas Cartas Constitucionais são considerados direitos fundamentais; e quando positivados na legislação civil são direitos de personalidade. ASCENSÃO, José de Oliveira. A dignidade da pessoa e o fundamento dos direitos humanos. *Revista da Faculdade de Direito da Universidade de São Paulo*, São Paulo, v. 103, p. 277-299, jan./dez., 2008. p. 278.

⁹⁰ Conferir, por todos, SARLET, Ingo Wolfgang. *A eficácia dos direitos fundamentais: uma teoria geral dos direitos fundamentais na perspectiva constitucional*. 10. ed. Porto Alegre: Livraria do Advogado, 2010. p. 79; HUNT, Lynn. *A invenção dos direitos humanos: uma história*. Tradução de Rosaura Eichenberg. São Paulo: Cia. das Letras, 2009. p. 113-145; COMPARATO, Fábio Konder. *A afirmação histórica dos direitos humanos*. 7. ed. São Paulo: Saraiva, 2010. p. 91-92; RECASÉNS SICHES, Luis. *Filosofía del derecho*. México: Porrúa, 2008. p. 1-19.

⁹¹ ZAGREBELSKY, Gustavo. *El derecho dúctil*. Ley, derechos y justicia. Tradução do italiano para o espanhol de Marina Gascón. Madri: Trotta, 1995. p. 15, tradução livre. No original: “a través de la red de valores y procedimientos comunicativos que es además la única visión no catastrófica de la política posible en nuestro tiempo.”

Embora não se possa deixar de considerar os impactos que as peculiaridades culturais acarretam para qualquer coletividade, a ponto de ser precipitada uma análise conjectural baseada na ideia de sociedade (ou “aldeia”⁹²) global, do ponto de vista dos direitos humanos, posições identitárias e individuais impõem a ponderação⁹³, notadamente para que sejam fixadas firmemente as bases do entrelaçamento entre direito público e direito privado⁹⁴.

A inteligência artificial afeta exatamente esse contexto, pois distorce os tradicionais *standards* de proteção historicamente construídos⁹⁵. Altera-se, essencialmente, a própria acepção que se tem do que é considerado “humano”⁹⁶. Tudo se reduz à matemática, à heurística, à predição e ao mapeamento dos interesses como condicionantes funcionais de estruturas de mercado, pois não se atingiu a já descrita singularidade tecnológica. As máquinas não são capazes de agir ou tomar decisões imbuídas de concepções morais.

Naturalmente, reconfigurar estruturas protetivas condizentes com o novo momento do desenvolvimento técnico-informacional. Os direitos humanos devem inspirar marcos regulatórios, propostas legislativas e, essencialmente, todo o acervo normativo que se pretenda instituir.

No Brasil, o que guiou a promulgação da Lei Geral de Proteção de Dados Pessoais foi justamente esse “núcleo duro” de parâmetros extraídos de uma compreensão mais ampla do direito à privacidade⁹⁷, previstos no texto legal como fundamentos, em seu artigo 2º, e que atuam como vetores axiológicos para os direitos (e deveres) descritos noutras passagens da lei e, também, para a atuação posterior do Estado, no exercício de seu poder regulatório. Destaca-se a redação do inciso VII do artigo 2º da LGPD: “A disciplina da proteção de dados pessoais tem como fundamentos: (...) VII - os direitos humanos, o livre desenvolvimento da personalidade, a dignidade e o exercício da cidadania pelas pessoas naturais”⁹⁸.

Na linha do que destaca Bárbara Dayana Brasil, “busca-se garantir que o novo ambiente de exercício dos direitos fundamentais, que é o plano digital, não fique à margem da estruturação de uma regulamentação jurídica apta a fornecer as mesmas garantias existentes no plano físico, com especial atenção para formas mais sofisticadas de violação que possam

⁹² Cf. McLuhan, H. Marshall; Fiore, Quentin. *Guerra e paz na aldeia global*. Tradução de Ivan Pedro de Martins. Rio de Janeiro: Record, 1971.

⁹³ HERRERA FLORES, Joaquín. *Teoria crítica dos direitos humanos: os direitos humanos como produtos culturais*. Tradução de Luciana Caplan. Rio de Janeiro: Lumen Juris, 2009. p. 97-98.

⁹⁴ SILVA, Jorge Pereira da. *Deveres do Estado de protecção de direitos fundamentais: fundamentação e estrutura das relações jusfundamentais triangulares*. 3. ed. Lisboa: Universidade Católica Editora, 2015. p. 354.

⁹⁵ GREENE, Kevin T.; PARK, Baekkwon; COLARESI, Michael. Machine learning human rights and wrongs: how the successes and failures of supervised learning algorithms can inform the debate about information effects. *Political Analysis*, Cambridge, v. 27, n. 2, p. 223-230, abr. 2019.

⁹⁶ Cf. COLEMAN, Flynn. *A human algorithm: how Artificial Intelligence is redefining who we are*. Berkeley: Counterpoint, 2019.

⁹⁷ TRSTENJAK, Verica. General report: The influence of human rights and basic rights in private law. In: TRSTENJAK, Verica; WEINGERL, Petra (ed.). *The influence of human rights in private law*. Cham: Springer, 2016. p. 48. “The right to privacy functions as a ‘window’ to the constitutional rights, which protects the private and personal sphere of an individual, e.g., dignity, free development of personality, religious freedom, privacy and data protection.”

⁹⁸ BRASIL, Bárbara Dayana. Os direitos humanos como fundamento da proteção de dados pessoais na Lei Geral de Proteção de Dados brasileira. In: CRAVO, Daniela Copetti; JOBIM, Eduardo; FALEIROS JÚNIOR, José Luiz de Moura (coord.). *Direito público e tecnologia*. Indaiatuba: Foco, 2022, p. 54.

ocorrer por força dos avanços tecnológicos”. Assim, a definição de categorias merecedoras de maior proteção, como a dos dados pessoais adjetivados como “sensíveis” (art. 5º, II, da LGPD) é evidência sólida dessa preocupação do legislador. Outra evidência disso é estruturação de revisões das decisões automatizadas (art. 20, da LGPD), que devem ser realizadas por agentes humanos. Se a proteção insuficiente não pode ser admitida, sob pena de flagrante violação ao citado fundamento da lei, deve-se estruturar mecanismos de controle que atuem como freios aos desideratos que afastem os humanos de sua essência⁹⁹.

Ainda, com o advento das novas tecnologias do século XXI e o surgimento do ambiente das mídias sociais, os países se veem involuntariamente envolvidos na sociedade digital global; aplicativos como *Facebook*, *WhatsApp* e *Twitter* estão presentes na vida da população e transformam o ambiente digital em espaço de interação social. Para os países da América Latina que lidam com problemas raciais estruturais, o ambiente virtual – ao contrário da ideia ilusória de ambiente “*colour-blind*”¹⁰⁰, onde características de raça seriam inexistentes e irrelevantes – constitui instituição de concretização de práticas racistas, discrepando da ideia de dissociação entre ambiente físico e virtual. Portando, novos debates sobre racismo devem surgir aliados ao ambiente digital, sobretudo estudos sobre o racismo algorítmico como ferramenta de reforço do racismo estrutural e ocultação de entenderes diversos do Ocidente¹⁰¹.

Como anota Romualdo Baptista dos Santos:

Lembremos que a dignidade humana é multidimensional porque se faz presente sempre e em qualquer situação em que se trate de um, de alguns, de muitos ou de todos os seres humanos. Então, o dano moral pode atingir uma pessoa em sua condição de pertencimento, pode atingir algumas pessoas, muitas pessoas ou todas as pessoas, dando azo às especificações: dano moral individual, dano moral coletivo e dano social. No dano moral individual, o prejuízo é sentido por uma pessoa ou por algumas pessoas individualizadas; no dano moral coletivo, o prejuízo alcança valores que são caros a uma coletividade; o dano social implica rebaixamento no nível de vida de toda a sociedade, envolvendo prejuízos morais e materiais. Há ainda o dano à humanidade, que configura atentado aos direitos humanos¹⁰².

Se os algoritmos não são capazes de “pensar”, mas são suficientemente avançados para tomar decisões “resultantes de uma combinação de *inputs* de programação não originária”¹⁰³, uma leitura ampliada do conceito de ‘responsabilidade’ poderá ao menos nortear a solução de

⁹⁹ MOYN, Samuel. *Not enough: human rights in an unequal world*. Cambridge: Harvard University Press, 2018. p. 220. O autor assevera: “Human rights will return to their defensible importance only as soon as humanity saves itself from its low ambitions. If it does, for the sake of local and global welfare, sufficiency and equality can again become powerful companions, both in our moral lives and in our political enterprises.”

¹⁰⁰ TURKLE, Sherry. *Life on the Screen*. Nova York: Simon & Schuster, 1995.

¹⁰¹ SILVA, Tarcizio. *Racismo Algorítmico em Plataformas Digitais: microagressões e discriminação em código*. *Anais do VI Simpósio Internacional Lavits - Assimetrias e (In)visibilidades: Vigilância, Gênero e Raça*, Salvador, BA, 26-28 junho 2019. Disponível em: https://lavits.org/vi-simposio-internacional-lavits-salvador-26_28-de-junho-2019/?lang=pt Acesso em: 18 mar. 2023.

¹⁰² SANTOS, Romualdo Baptista dos. Responsabilidade civil pela prática de racismo: danos morais individuais e coletivos. *Revista Fórum de Direito Civil*, Belo Horizonte, ano 10, n. 27, p. 37-56, maio/ago. 2021, p. 46.

¹⁰³ FERREIRA, Ana Elisabete. Responsabilidade civil extracontratual por danos causados por robôs autônomos: breves reflexões. *Revista Portuguesa do Dano Corporal*, Lisboa, n. 27, 39-63, dez, 2016, p. 44.

problemas como “a opacidade decisória, a falta de explicação quanto aos critérios utilizados e a herança de *inputs* viciados, enviesados e preconceituosos, o que culmina na produção de discriminações injustificadas”¹⁰⁴. É o que se espera com o fomento à *accountability*¹⁰⁵.

Como foi dito na introdução, alguns documentos importantíssimos, como o *Artificial Intelligence Act* europeu de 2021¹⁰⁶ (2021 EU AIA) e o recentíssimo *Algorithmic Accountability Act* norte-americano de 2022¹⁰⁷ (2022 US AAA), que atualizou a versão anterior, de 2019¹⁰⁸, evitam a discussão terminológica sobre o alcance semântico do termo “inteligência”, preferindo se reportar a “sistemas decisoriais automatizados”¹⁰⁹ (*Automated Decision Systems*, ou ADS’s) ao se referir à necessidade de que seja definido um regime de responsabilidade civil aplicável a eventos danosos propiciados por tais sistemas e até mesmo à estruturação de parâmetros éticos para o desenvolvimento de algoritmos. Segundo abalizada doutrina, os documentos citados possuem qualidades que podem servir para mútua inspiração, denotando a importância da adequada assimilação semântica (além de outros temas) para a evolução das discussões¹¹⁰.

No Brasil, os Projetos de Lei n°s 5.051/2019, 21/2020 e 872/2021 visam regulamentar o tema em linhas gerais (e não apenas para o contexto dos carros autônomos), priorizando a delimitação de um sistema de responsabilização baseado na anacrônica teoria da culpa, que simplesmente não faz sentido para tutelar matéria tão complexa.

Todavia, em fevereiro de 2022, foi instituída, pelo Senado Federal, a elogiável “Comissão de Juristas responsável por subsidiar elaboração de substitutivo sobre IA” (CJSUBIA)¹¹¹, que realizou diversas reuniões e audiências públicas nos trabalhos de elaboração do substitutivo, que foi apresentado em dezembro de 2022 e inspirou o texto final do Projeto de Lei n° 2.338/2023, apresentado pelo Senador Rodrigo Pacheco. Dentre suas diversas previsões,

¹⁰⁴ MEDON, Filipe. *Inteligência artificial e responsabilidade civil: autonomia, riscos e solidariedade*. 2. ed. Salvador: Juspodivm, 2021, p. 525.

¹⁰⁵ Explica Matheus Schwede que “as preocupações que vieram à tona, envolvendo governos e diversas organizações, têm movimentado esforços para propor estruturas políticas para tratar dos temas de *fairness* (justiça), *accountability*, *transparency* (transparência) e outros tópicos relacionados (...) [que] realçaram a importância de se trabalhar *accountability* para os modelos de inteligência artificial, especialmente no que tange a alcançar um design, desenvolvimento e uso responsável de inteligências artificiais”. SCHWEDE, Matheus. *Antes. Inteligência artificial tendenciosa: discriminação pelo viés algorítmico*. Londrina: Thoth, 2023, p. 108.

¹⁰⁶ EUROPA. European Commission. *Artificial Intelligence Act*. 2021/0106(COD), abr. 2021. Disponível em: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206>. Acesso em: 18 mar. 2023.

¹⁰⁷ ESTADOS UNIDOS DA AMÉRICA. House of Representatives. House Resolution No. 6580, Feb. 3, 2022. *Algorithmic Accountability Act of 2022*. Disponível em: <https://www.congress.gov/bill/117th-congress/house-bill/6580/text> Acesso em: 18 mar. 2023.

¹⁰⁸ ESTADOS UNIDOS DA AMÉRICA. House of Representatives. House Resolution No. 2231, Apr. 10, 2019. *Algorithmic Accountability Act of 2019*. Disponível em: <https://www.congress.gov/116/bills/hr2231/BILLS-116hr2231ih.pdf> Acesso em: 18 mar. 2023.

¹⁰⁹ Cf. SELBST, Andrew. An institutional view of algorithmic impact assessments. *Harvard Journal of Law & Technology*, Cambridge, v. 35, 2021. Disponível em: <https://ssrn.com/abstract=3867634> Acesso em: 18 mar. 2023.

¹¹⁰ MÖKANDER, Jakob; JUNEJA, Pratham; WATSON, David S.; FLORIDI, Luciano. The US Algorithmic Accountability Act of 2022 vs. The EU Artificial Intelligence Act: what can they learn from each other? *Minds and Machines*, Cham: Springer, v. 22, p. 1-9, jun. 2022. Disponível em: <https://doi.org/10.1007/s11023-022-09612-y> Acesso em: 18 mar. 2023.

¹¹¹ BRASIL. Senado Federal. Atividade Legislativa. *Comissão de Juristas responsável por subsidiar elaboração de substitutivo sobre IA (CJSUBIA)*. Disponível em: <https://legis.senado.leg.br/comissoes/comissao?codcol=2504> Acesso em: 18 jun. 2023.

destacam-se os princípios elencados nos incisos X e XI do artigo 3º, cujas previsões impõem que “o desenvolvimento, a implementação e o uso de sistemas de inteligência artificial observarão a boa-fé e os seguintes princípios: (...) X – prestação de contas, responsabilização e reparação integral de danos; XI – prevenção, precaução e mitigação de riscos sistêmicos derivados de usos intencionais ou não intencionais e de efeitos não previstos de sistemas de inteligência artificial”¹¹². Além disso, o Capítulo III do PL é integralmente dedicado à “categorização dos riscos”, que demandam avaliação preliminar (art. 13), permitindo classificações nas categorias de ‘risco excessivo’ (arts. 14 a 16) e ‘alto risco’ (arts. 17 e 18), impondo-se a ‘governança algorítmica’ (Capítulo IV, arts. 19 a 21) e a ‘avaliação de impacto algorítmico’ (arts. 22 a 26). E, em matéria de responsabilidade civil, ainda se define uma cláusula geral de reparação do dano independentemente do grau de risco (art. 27), excetuando-se o contexto das relações de consumo (art. 29)¹¹³.

Primando pela proteção da confiança¹¹⁴, a discussão sobre *accountability* contempla aspectos éticos que passam a ser reinterpretados para a aplicação de algoritmos de inteligência artificial no campo de investigação que concerne à necessidade de regulação específica. Ao invés de normatizar “como” devem ser criados, aplicados e fiscalizados os algoritmos, esse modelo mais aberto cria nichos específicos nos quais faz mais sentido estabelecer *guidelines* mais abstratas, com grau de generalização suficiente para orientar o desenvolvimento tecnológico sem ceifá-lo da ambientação profícua ao seu livre florescimento.

De um lado, busca-se tutelar aquilo que, embora seja uma criação humana, instiga e desafia o próprio homem e seus limites; de outro, surge a *accountability* como paradigma de reforço à proliferação de verdadeira cultura baseada no mapeamento de riscos e na conformação das estruturas de pesquisa, produção e desenvolvimento de novas tecnologias, com vistas ao aumento do rol de contingências previsíveis e à conversão de perigos em riscos, ou do incerto em plausível para que se possa aliviar esse dilema¹¹⁵.

¹¹² BRASIL. Senado Federal. Projeto de Lei nº 2.338, de 2023. *Dispõe sobre o uso da Inteligência Artificial*. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/157233> Acesso em: 18 jun. 2023.

¹¹³ BRASIL. Senado Federal. Projeto de Lei nº 2.338, de 2023. *Dispõe sobre o uso da Inteligência Artificial*. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/157233> Acesso em: 18 jun. 2023.

¹¹⁴ FRADA, Manuel A. Carneiro da. *Teoria da confiança e responsabilidade civil*. Coimbra: Almedina, 2004, p. 17. Anota: “Não existe nenhuma definição legal de confiança a que possa socorrer-se e escasseiam referências normativas explícitas a propósito. O seu conceito apresenta-se fortemente indeterminado pela pluralidade ou vaguidade de empregos comuns que alberga, tornando difícil traçar com ele as fronteiras de uma investigação jurídica. Tanto mais que transporta uma certa ambiguidade de princípio por se poder referir, tanto à causa, como aos efeitos de uma regulação jurídica. É a falta de consciência desta realidade que está na raiz de uma certa evanescência da confiança no discurso jurídico e se apresenta – antecipe-se – responsável pelas dificuldades de que se não logrou desembaraçar-se a reflexão dogmática a seu respeito.”

¹¹⁵ Segundo Comandé, “il costo dell’osservazione sistematica dei comportamenti pericolosi degli utenti può anche diventare sufficientemente basso da rendere più efficiente per i produttori l’adozione di misure precauzionali attraverso il ridisegno del prodotto”. COMANDÉ, Giovanni. *Intelligenza Artificiale e responsabilità tra liability e accountability: il carattere trasformativo dell’IA e il problema della responsabilità*. In: NUZZO, Antonio; OLIVIERI, Gustavo (a cura di) *Analisi giuridica dell’Economia*. Studi e discussioni sul diritto dell’impresa. Bologna: Il Mulino, 2019, v. 1, p. 185.

5. CONCLUSÃO

A ‘singularidade tecnológica’ ainda não foi atingida e não se tem sistemas de inteligência artificial efetivamente inteligentes, com funcionamento idêntico ao dos cérebros humanos, ainda que o fascínio pela técnica e a rápida evolução computacional tornem essa perspectiva cada vez mais convidativa e instigante. De fato, ao contrário que qualquer empolgação distópica possa sugerir, as estruturas algorítmicas ainda se baseiam na heurística computacional e em processos decisoriais baseados no processamento de dados e nas inferências produzidas a partir do que a doutrina designa como “predição por números”, inclusive em modelos mais complexos como os que dizem respeito às técnicas de aprendizado profundo (*deep learning*).

Por essa razão, a resposta à indagação lançada no tema-problema desse artigo perpassa pela compreensão da distinção entre os conceitos de “preconceito” e “viés”, que, como visto, não se confundem e, sendo distintos, confirmam a hipótese de pesquisa com uma afirmação negativa, qual seja: não há, no atual estado da técnica, algoritmos racistas.

Logo, sistemas de inteligência artificial não podem ser adjetivados como tal. Por outro lado, não se pode desconsiderar o impacto nefasto que vieses algorítmicos podem ter se forem alimentados por conjuntos de dados enviesados, pois situações danosas já ocorreram, a exemplo dos casos envolvendo o viés racial indicado na falha algorítmica pela propensão ao atropelamento de pessoas negras com a tecnologia *LiDAR* em carros autônomos. E é a partir disso que surge o imprescindível debate sobre a parametrização ética para o desenvolvimento algorítmico que, alinhando-se ao propósito de sistematizar deveres norteadores do desenvolvimento e do bom implemento desses sistemas heurísticos, podem conduzir à responsabilização daqueles que deixam de observar tais parâmetros, agindo omissivamente.

Por todo o exposto, a parametrização realizada por humanos ainda é o fator preponderante para que um algoritmo se torne enviesado e falhe, gerando danos que se convolvam em ilícito indenizável. Sua falta de autonomia, contudo, afasta qualquer possibilidade de adjetivação de suas inferências como preconceitos, ainda que o ser humano possa ser preconceituoso e transmitir – mesmo que sutilmente – seus preconceitos ao conjunto de dados que será processado pelo algoritmo, por indevida parametrização; já o algoritmo, dependente que é da heurística, ainda não pode sê-lo isoladamente.

Há longo percurso doutrinário em torno dessa matéria, o que conduz à necessidade de reconhecimento dos perigos da utilização abstrusa de sistemas de inteligência artificial sem a devida parametrização ética. Esse é o debate fundamental que deve permear os estudos sobre aspectos éticos para o desenvolvimento algorítmico, com repercussões sobre a disciplina da responsabilidade civil.

E, nesse exato sentido, debates regulatórios, como o que consta do *AI Act* europeu de 2021, se mostram propensos ao acolhimento de uma abordagem baseada em riscos (*risk-based approach*), que estratifica os usos permitidos, proibidos ou limitados de sistemas de inteligência artificial, a depender da reprovabilidade de eventual consequência que possam vir a causar. E

também parece ser essa a propensão regulatória brasileira, que, após elogiável trabalho de importante Comissão de Juristas, culminou na apresentação do Projeto de Lei nº 2.338/2023. Esse parece ser o melhor caminho a trilhar, inclusive para que se permita uma abordagem analítica da teoria do risco para inspirar estudos em matéria de responsabilidade civil objetiva para falhas danosas de sistemas de IA enviesados.

REFERÊNCIAS

ALMEIDA, Silvio Luiz de. *Racismo estrutural*. São Paulo: Pólen, 2019.

ARMSTRONG, Edwin H. Operating field of the audion. *Electrical World*, Nova York, v. 54, n. 1, p. 1149-1152, jul./dez. 1914. Disponível em: <https://bit.ly/2NLe0Qf>. Acesso em: 18 mar. 2023.

ASCENSÃO, José de Oliveira. A dignidade da pessoa e o fundamento dos direitos humanos. *Revista da Faculdade de Direito da Universidade de São Paulo*, São Paulo, v. 103, jan./dez., 2008.

BARBOSA, Mafalda Miranda. Inteligência artificial, e-persons e direito: desafios e perspectivas. *Revista Jurídica Luso-Brasileira*, Lisboa, ano 3, n. 6, 2017.

BÀRCENA, Alicia. *Apelo para reforçar o compromisso com os direitos humanos das populações afrodescendentes e redobrar esforços no combate ao racismo e à discriminação*. 2020. Disponível em: <https://www.cepal.org/pt-br/noticias/apelo-reforcar-o-compromisso-os-direitos-humanos-populacoes-afrodescendentes-redobrar>. Acesso em: 18 mar. 2023.

BORGES, Gustavo Silveira; LIMA, Fernanda da Silva. Publicidade e racismo reverso: o que uma campanha publicitária tem a revelar sobre o racismo no Brasil. *Revista de Direito do Consumidor*, São Paulo, v. 123, p. 37-76, 2019.

BOSTROM, Nick. *Superintelligence: paths, dangers, strategies*. Oxford: Oxford University Press, 2014.

BRASIL. Senado Federal. Atividade Legislativa. *Comissão de Juristas responsável por subsidiar elaboração de substitutivo sobre IA (CJSUBIA)*. Disponível em: <https://legis.senado.leg.br/comissoes/comissao?codcol=2504> Acesso em: 18 jun. 2023.

BRASIL. Senado Federal. Projeto de Lei nº 2.338, de 2023. *Dispõe sobre o uso da Inteligência Artificial*. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/material/157233> Acesso em: 18 jun. 2023.

BRASIL, Bárbara Dayana. Os direitos humanos como fundamento da proteção de dados pessoais na Lei Geral de Proteção de Dados brasileira. In: CRAVO, Daniela Copetti; JOBIM, Eduardo; FALEIROS JÚNIOR, José Luiz de Moura (coord.). *Direito público e tecnologia*. Indaiatuba: Foco, 2022.

CASTELLS, Manuel. *The rise of the network society*. The information age: economy, society, and culture 2. ed. Oxford/West Sussex: Wiley-Blackwell, 2010, v. 1.

CAVE, Stephen; DIHAL, Kanta. The whiteness of AI. *Philosophy & Technology*, Londres, v. 33, 2020.

CERUZZI, Paul E. *A history of modern computing*. 2. ed. Cambridge: The MIT Press, 2003.

CHANDER, Anupam. The racist algorithm? *Michigan Law Review*, Ann Arbor, v. 115, n. 6, p.

1023-1045, 2017.

CICUREL, Ronald; NICOLELIS, Miguel. *O cérebro relativístico: como ele funciona e por que ele não pode ser simulado por uma máquina de Turing*. Natal/Montreux/Durham/São Paulo: Kios Press, 2015.

COLEMAN, Flynn. *A human algorithm: how Artificial Intelligence is redefining who we are*. Berkeley: Counterpoint, 2019.

COMANDÉ, Giovanni. Intelligenza Artificiale e responsabilità tra liability e accountability: il carattere trasformativo dell'IA e il problema della responsabilità. In: NUZZO, Antonio; OLIVIERI, Gustavo (a cura di) *Analisi giuridica dell'Economia*. Studi e discussioni sul diritto dell'impresa. Bologna: Il Mulino, 2019, v. 1.

COMISSIÓN ECONÓMICA PARA AMÉRICA LATINA Y EL CARIBE (CEPAL). *Panorama Social de América Latina*, 2016. Santiago, 2017.

COMPARATO, Fábio Konder. *A afirmação histórica dos direitos humanos*. 7. ed. São Paulo: Saraiva, 2010.

CORRÊA, Bianca Kremer Nogueira. *Direito e tecnologia em perspectiva ameíricana: autonomia, algoritmos e vieses raciais*. 2021. 299p. Tese (Doutorado em Direito) – Faculdade de Direito, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2021.

CUTHBERTSON, Anthony. Self-driving cars more likely to drive into black people, study claims. *The Independent*, 06 mar. 2019. Disponível em: <https://www.independent.co.uk/life-style/gadgets-and-tech/news/self-driving-car-crash-racial-bias-black-people-study-a8810031.html>. Acesso em: 18 mar. 2023.

DOMINGOS, Pedro. *The master algorithm: how the quest for the ultimate learning machine will remake our world*. Nova York: Basic Books, 2015.

DONEDA, Danilo Cesar Maganhoto; MENDES, Laura Schertel; SOUZA, Carlos Affonso Pereira de; ANDRADE, Norberto Nuno Gomes de. 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.

DUSSEL, Enrique. Europa, modernidade e eurocentrismo. In: LANDER, Edgardo (org.). *A colonialidade do saber: eurocentrismo e ciências sociais, perspectivas latinoamericanas*. São Paulo: CLACSO, 2005.

EBERHARDT, Jennifer. *Biased: uncovering the hidden prejudice that shapes what we see, think, and do*. Nova York: Viking, 2019.

EHRHARDT JÚNIOR, Marcos; SILVA, Gabriela Buarque Pereira. Pessoa e sujeito de direito: reflexões sobre a proposta europeia de personalidade jurídica eletrônica. *Revista Brasileira de Direito Civil*, Belo Horizonte, v. 23, n. 1, p. 57-79, jan./mar. 2020.

ESCOBAR, Arturo. *El giro decolonial: reflexiones para una diversidad epistémica más allá del capitalismo global*. 2007.

ESTADOS UNIDOS DA AMÉRICA. House of Representatives. House Resolution No. 2231, Apr. 10, 2019. *Algorithmic Accountability Act of 2019*. Disponível em: <https://www.congress.gov/116/bills/hr2231/BILLS-116hr2231ih.pdf> Acesso em: 18 mar. 2023.

ESTADOS UNIDOS DA AMÉRICA. House of Representatives. House Resolution No. 6580, Feb. 3, 2022. *Algorithmic Accountability Act of 2022*. Disponível em: <https://www.congress.gov/bill/117th-congress/house-bill/6580/text> Acesso em: 18 mar. 2023.

- EUROPA. European Commission. *Artificial Intelligence Act*. 2021/0106(COD), abr. 2021. Disponível em: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206>. Acesso em: 18 mar. 2023.
- FALEIROS JÚNIOR, José Luiz de Moura. Breves reflexões sobre os impactos jurídicos do algoritmo GPT-3. In: BARBOSA, Mafalda Miranda et al (coord.). *Direito digital e inteligência artificial: diálogos entre Brasil e Europa*. Indaiatuba: Foco, 2021.
- FALEIROS JÚNIOR, José Luiz de Moura. Discriminação por algoritmos de Inteligência Artificial: a responsabilidade civil, os vieses e o exemplo das tecnologias baseadas em luminância. *Revista de Direito da Responsabilidade*, Coimbra, ano 2, 2020.
- FALEIROS JÚNIOR, José Luiz de Moura; MENKE, Fabiano. "Teilrechtsfähigkeit": uma proposta alemã para a responsabilização civil na IA. *Migalhas de Responsabilidade Civil*, 06 ago. 2020. Disponível em: <https://s.migalhas.com.br/S/8AF9D>. Acesso em: 18 mar. 2023.
- FANON, Frantz. *Os condenados da terra*. Tradução de Enilce Albergaria Rocha e Lucy Magalhães. Juiz de Fora (MG): Editora da UFJF, 2005.
- FERREIRA, Ana Elisabete. Responsabilidade civil extracontratual por danos causados por robôs autônomos: breves reflexões. *Revista Portuguesa do Dano Corporal*, Lisboa, n. 27, 39-63, dez, 2016.
- FLASIŃSKI, Mariusz. *Introduction to Artificial Intelligence*. Cham: Springer, 2016.
- FRADA, Manuel A. Carneiro da. *Teoria da confiança e responsabilidade civil*. Coimbra: Almedina, 2004.
- GANASCIA, Jean-Gabriel. *Le mythe de la singularité: faut-il craindre l'intelligence artificielle?* Paris: Éditions du Seuil, 2017.
- GARDNER, Howard. *Frames of mind: the theory of multiple intelligences*. Nova York: Basic Books, 2011.
- GOETTENAUER, Carlos Eduardo. Algoritmos, inteligência artificial, mercados. Desafios ao arcabouço jurídico. In: FRAZÃO, Ana; CARVALHO, Angelo Gamba Prata de (coord.). *Empresa, mercado e tecnologia*. Belo Horizonte: Fórum, 2019.
- GREENE, Kevin T.; PARK, Baekkwon; COLARESI, Michael. Machine learning human rights and wrongs: how the successes and failures of supervised learning algorithms can inform the debate about information effects. *Political Analysis*, Cambridge, v. 27, n. 2, abr. 2019.
- GROSFUGUEL, Ramón. Para descolonizar os estudos de economia política e os estudos pós-coloniais: transmodernidade, pensamento de fronteira e colonialidade global. *Revista Crítica de Ciências Sociais*, Coimbra, n. 80, 2008.
- HARNAD, Stevan. The Annotation Game: on Turing (1950), on computing, machinery and intelligence. In: EPSTEIN, Robert; PETERS, Grace (Ed.). *Parsing the Turing Test: philosophical and methodological issues in the quest for the thinking computer*. Cham: Springer, 2008.
- HERNÁNDEZ, Tanya Katerí; SOUZA, Arivaldo Santos de; FONSECA, Luciana Carvalho. *Subordinação racial no Brasil e na América Latina: o papel do Estado, o Direito Costumeiro e a Nova Resposta dos Direitos Cíveis*. EDUFBA, 2017.
- HERRERA FLORES, Joaquín. *Teoria crítica dos direitos humanos: os direitos humanos como produtos culturais*. Tradução de Luciana Caplan. Rio de Janeiro: Lumen Juris, 2009.

HOFFMANN-RIEM, Wolfgang. Inteligência artificial como oportunidade para a regulação jurídica. Tradução de Luiz Sander; revisão de Laura Schertel Mendes. *Revista Direito Público*, Porto Alegre, v. 16, n. 90, p. 11-38, nov./dez. 2019.

HUNT, Lynn. *A invenção dos direitos humanos: uma história*. Tradução de Rosaura Eichenberg. São Paulo: Cia. das Letras, 2009.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Desigualdades Sociais por Cor ou Raça no Brasil*. Estudos e Pesquisas: Informação Demográfica e Socioeconômica, n. 41, [S.D]. Disponível em https://biblioteca.ibge.gov.br/visualizacao/livros/liv101681_informativo.pdf Acesso em: 18 mar. 2023.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Pesquisa Nacional por Amostra de Domicílios (Pnad) Contínua*. 2019. Disponível em: <https://sidra.ibge.gov.br/tabela/6403> Acesso em: 18 mar. 2023.

INSTITUTO DE PESQUISA ECONÔMICA APLICADA. Atlas da violência. Brasília/Rio de Janeiro/São Paulo: IPEA, *Fórum Brasileiro de Segurança Pública*, 2019. v. 11, 2020.

KAHNEMAN, Daniel; TVERSKY, Amos. On the psychology of prediction. In: KAHNEMAN, Daniel; SLOVIC, Paul; TVERSKY, Amos (ed.). *Judgement under uncertainty: heuristics and biases*. Cambridge: Cambridge University Press, 1982.

KAPLAN, Jerry. *Humans need not apply: a guide to wealth and work in the Age of Artificial Intelligence*. New Haven: Yale University Press, 2015.

KELLEHER, John D. *Deep learning*. Cambridge: The MIT Press, 2019.

KELLEHER, John D.; MAC NAMEE, Brian; D'ARCY, Aiofe. *Fundamentals of machine learning for predictive data analytics: algorithms, worked examples, and case studies*. Cambridge: The MIT Press, 2015.

KITTLER, Friedrich. There is no software. *CTHEORY.net*. 18 out. 1995. Disponível em: <http://www.ctheory.net/articles.aspx?id=74>. Acesso em: 18 mar. 2023.

KURZWEIL, Ray. *The age of spiritual machines: when computers exceed human intelligence*. Nova York: Viking, 1999.

MADSBJERG, Christian. *Sensemaking: the power of the humanities in the age of the algorithm*. Nova York: Hachette, 2017. *E-book*.

MALDONADO-TORRES, Nelson. Sobre la colonialidade del ser: contribuciones AL desarrollo de un concepto. In: CASTRO-GOMEZ, Santiago; GROSFUGUEL, Ramón (org.). *El giro decolonial: reflexiones para una diversidad epistémica más Allá del capitalismo global*. Bogotá: Universidad Javeriana-Instituto Pensar/Universidad Central-IESCO/Siglo del Hombre Editores, 2007

MBEMBE, Achille. *Crítica da razão negra*. São Paulo: N-1, 2018.

McCARTHY, John. *Artificial Intelligence (complete) – Thinking Allowed* [Entrevistador: Jeffrey Mishlove]. YouTube (video). 3 nov. 2011. Disponível em: <https://www.youtube.com/watch?v=Ozipf13jRr4>. Acesso em: 18 jun. 2023.

McLUHAN, H. Marshall; FIORE, Quentin. *Guerra e paz na aldeia global*. Tradução de Ivan Pedro de Martins. Rio de Janeiro: Record, 1971.

MEDON, Filipe. *Inteligência artificial e responsabilidade civil: autonomia, riscos e solidariedade*. 2. ed. Salvador: Juspodivm, 2021.

- MELO, Miguel Ângelo Silva de; BANDEIRA, João Adolfo Ribeiro. Crítica aos estereótipos e ideias racistas no Brasil sob o prisma dos estudos Pós-coloniais. *Revista Direito e Práxis*, v. 7, n. 15, 2016.
- MENDES, Laura Schertel; MATTIUZZO, Marcela. Discriminação algorítmica: conceito, fundamento legal e tipologia. *Revista Direito Público*, Porto Alegre, v. 16, n. 90, p. 39-64, nov./dez. 2019.
- MENDES, Maria Manuela. Raça e racismo: controvérsias e ambiguidades / Race and racism: controversies and ambiguities. *Vivência: Revista de Antropologia*, v. 1, n. 39, 8 maio 2012.
- MICHALEWICZ, Zbigniew; FOGEL, David B. *How to solve it: Modern heuristics*. Cham: Springer, 2000.
- MILAGRES, Marcelo de Oliveira. A robótica e as discussões sobre a personalidade eletrônica. In: EHRHARDT JÚNIOR, Marcos; CATALAN, Marcos; MALHEIROS, Pablo (coord.). *Direito civil e tecnologia*. Belo Horizonte: Fórum, 2020.
- MILES, Robert; BROWN, Malcolm. *Racism*. Londres: Routledge, 2003.
- MÖKANDER, Jakob; JUNEJA, Prathm; WATSON, David S.; FLORIDI, Luciano. The US Algorithmic Accountability Act of 2022 vs. The EU Artificial Intelligence Act: what can they learn from each other? *Minds and Machines*, Cham: Springer, v. 22, p. 1-9, jun. 2022. Disponível em: <https://doi.org/10.1007/s11023-022-09612-y> Acesso em: 18 mar. 2023.
- MOYN, Samuel. *Not enough: human rights in an unequal world*. Cambridge: Harvard University Press, 2018.
- MULHOLLAND, Caitlin. Responsabilidade civil e processos decisórios autônomos em sistemas de Inteligência Artificial (IA): autonomia, imputabilidade e responsabilidade. In: FRAZÃO, Ana; MULHOLLAND, Caitlin (coord.). *Inteligência Artificial e Direito: ética, regulação e responsabilidade*. São Paulo: Thomson Reuters Brasil, 2019.
- MUNTADAS, Borja. *Inmediatez*. Capitalismo y vidas aceleradas. Barcelona: Chiado Editorial, 2016.
- NOBLE, Safiya Umoja. *Algorithms of oppression: how search engines reinforce racism*. Nova York: NYU Press, 2018.
- OLIVEIRA, Luiz Fernandes de. *Educação e militância decolonial*. Rio de Janeiro: Editora Selo Novo, 2018.
- ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. *Artificial intelligence Principles*. Recommendation on the Council on Artificial Intelligence. 2019. Disponível em: <https://oecd.ai/en/ai-principles> Acesso em: 18 mar. 2023.
- PARLAMENTO EUROPEU. *Resolução de 16 de fevereiro de 2017*. Disposições de Direito Civil sobre Robótica. Disponível em: https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_PT.html. Acesso em: 18 mar. 2023.
- PEPPET, Scott R. Regulating the Internet of Things: first steps toward managing discrimination, privacy, security, and consent. *Texas Law Review*, Austin, v. 93, 2014.
- QUIJANO, Aníbal. Colonialidade do poder e classificação social. *Epistemologias do Sul*. São Paulo: Cortez, 2010.
- RECASÉNS SICHES, Luis. *Filosofía del derecho*. México: Porrúa, 2008.

SAMUEL, Sigal. A new study finds a potential risk with self-driving cars: failure to detect dark-skinned pedestrians. *Vox*, 06 mar. 2019. Disponível em: <https://www.vox.com/future-perfect/2019/3/5/18251924/self-driving-car-racial-bias-study-autonomous-vehicle-dark-skin>. Acesso em: 18 mar. 2023.

SÁNCHEZ, Margarita; BRYAN Maurice. *Afro-descendants, discrimination and economic exclusion in Latin America*. London: Minority Rights Group International, maio 2003.

SANTOS, Romualdo Baptista dos. Responsabilidade civil pela prática de racismo: danos morais individuais e coletivos. *Revista Fórum de Direito Civil*, Belo Horizonte, ano 10, n. 27, p. 37-56, maio/ago. 2021.

SARLET, Ingo Wolfgang. *A eficácia dos direitos fundamentais: uma teoria geral dos direitos fundamentais na perspectiva constitucional*. 10. ed. Porto Alegre: Livraria do Advogado, 2010.

SCHIRMER, Jan-Erik. Artificial Intelligence and legal personality. "Teilrechtsfähigkeit": A partial legal status made in Germany. In: WISCHMEYER, Thomas; RADEMACHER, Timo (ed.). *Regulating Artificial Intelligence*. Cham: Springer, 2020.

SCHWAB, Klaus. *A quarta revolução industrial*. Tradução de Daniel Moreira Miranda. São Paulo: Edipro, 2016.

SCHWEDE, Matheus Antes. *Inteligência artificial tendenciosa: discriminação pelo viés algorítmico*. Londrina: Thoth, 2023.

SELBST, Andrew. An institutional view of algorithmic impact assessments. *Harvard Journal of Law & Technology*, Cambridge, v. 35, 2021. Disponível em: <https://ssrn.com/abstract=3867634>. Acesso em: 18 mar. 2023.

SILVA, Jorge Pereira da. *Deveres do Estado de proteção de direitos fundamentais: fundamentação e estrutura das relações jusfundamentais triangulares*. 3. ed. Lisboa: Universidade Católica Editora, 2015.

SILVA, Tarcízio. Racismo Algorítmico em Plataformas Digitais: microagressões e discriminação em código. *Anais do VI Simpósio Internacional Lavits - Assimetrias e (In)visibilidades: Vigilância, Gênero e Raça*, Salvador, BA, 26-28 junho 2019. Disponível em: https://lavits.org/vi-simposio-internacional-lavits-salvador-26_28-de-junho-2019/?lang=pt. Acesso em: 18 mar. 2023.

SOLUM, Lawrence. Legal personhood for Artificial Intelligences. *North Carolina Law Review*, Chapel Hill, v. 70, n. 4, 1992.

SUNSTEIN, Cass R. *Valuing life: humanizing the regulatory state*. Chicago: The University of Chicago Press, 2014.

TOMASECIVIUS FILHO, Eduardo. Inteligência Artificial e direitos da personalidade: uma contradição em termos? *Revista da Faculdade de Direito da Universidade de São Paulo*, São Paulo, v. 113, jan./dez. 2018.

TRSTENJAK, Verica. General report: The influence of human rights and basic rights in private law. In: TRSTENJAK, Verica; WEINGERL, Petra (ed.). *The influence of human rights in private law*. Cham: Springer, 2016.

TURING, Alan M. Computing machinery and intelligence. *Mind*, Oxford, n. 236, p. 433-460, out. 1950. Disponível em: <https://dx.doi.org/10.1093/mind/LIX.236.433>. Acesso em: 18 mar. 2023.

TURING, Alan M. On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, Londres, v. 42, n. 1, nov. 1936.

TURKLE, Sherry. *Life on the Screen*. Nova York: Simon & Schuster, 1995.

VALLAS, Steven Peter. *Work*. Cambridge: Polity, 2012.

VAN DIJK, Jan. *The network society*. 2. ed. Londres: Sage, 2006.

VARANASI, Lakshmi. AI models like ChatGPT and GPT-4 are acing everything from the bar exam to AP Biology. Here's a list of difficult exams both AI versions have passed. *Business Insider*, 25 jun. 2023. Disponível em: <https://www.businessinsider.com/list-here-are-the-exams-chatgpt-has-passed-so-far-2023-1> Acesso em: 26 jun. 2023.

VINGE, Vernor. The coming technological singularity: How to survive in the post-human era. In: Interdisciplinary Science and Engineering in the Era of Cyberspace. *NASA John H. Glenn Research Center at Lewis Field*, Cleveland, 1993. Disponível em: <https://ntrs.nasa.gov/search.jsp?R=19940022856>. Acesso em: 18 mar. 2023.

WARWICK, Kevin. The disappearing human-machine divide. In: ROMPORTL, Jan; ZACKOVA, Eva; KELEMEN, Jozef (ed.). *Beyond Artificial Intelligence: the disappearing human-machine divide*. Cham: Springer, 2015.

ZAGREBELSKY, Gustavo. *El derecho dúctil*. Ley, derechos y justicia. Tradução do italiano para o espanhol de Marina Gascón. Madri: Trotta, 1995.

ZONINSEIN, Jonas. The economic case for combating racial and ethnic exclusion in Latin America and the Caribbean Countries. In: BUVINIC, Mayra; MAZZA, Jacqueline; DEUTSCH, Ruthanne. (ed.). *Towards a shared vision of development*. Washington, DC: InterAmerican Development Bank, 2001.

Recebido: 17.09.2022

Aprovado: 26.06.2023

Como citar: BORGES, Gustavo Silveira; FALEIROS JÚNIOR, José Luiz de Moura. Viés racial em sistemas de inteligência artificial: desafios para a responsabilidade civil e a necessidade de parametrização do risco. **Revista IBERC**, Belo Horizonte, v. 6, n. 2, p. 100-128, maio/ago. 2023.

