

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
FACULDADE DE DIREITO  
DEPARTAMENTO DE DIREITO ECONÔMICO E DO TRABALHO

ALLAN PEREIRA DE CARVALHO

**VIÉS ALGORÍTMICO E DISCRIMINAÇÃO:  
Possíveis soluções regulatórias para o Brasil**

Porto Alegre  
2020

ALLAN PEREIRA DE CARVALHO

**VIÉS ALGORÍTMICO E DISCRIMINAÇÃO:  
Possíveis soluções regulatórias para o Brasil**

Monografia apresentada a título de trabalho de conclusão de curso como requisito parcial para obtenção do grau de Bacharel no curso de graduação da Faculdade de Direito da Universidade Federal do Rio Grande do Sul.

Orientadora: Prof. Dra. Kelly Lissandra Bruch.

Porto Alegre  
2020

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

ALLAN PEREIRA DE CARVALHO

**VIÉS ALGORÍTMICO E DISCRIMINAÇÃO:  
Possíveis soluções regulatórias para o Brasil**

Monografia apresentada a título de trabalho de conclusão de curso como requisito parcial para obtenção do grau de Bacharel no curso de graduação da Faculdade de Direito da Universidade Federal do Rio Grande do Sul.

Orientadora: Prof. Dra. Kelly Lissandra Bruch.

**BANCA EXAMINADORA**

---

Orientadora – Prof. Dra. Kelly Lissandra Bruch

---

Prof. Dra. Maria Cláudia Mércio Cachapuz

---

Prof. Me. Maurício Brum Esteves

*aos meus avós, pais, Isabella e Erick*

## AGRADECIMENTOS

Agradeço à Universidade Federal do Rio Grande do Sul, na figura dos seus alunos, professores, servidores e terceirizados, que tanto contribuíram para meu desenvolvimento pessoal e profissional. É um privilégio ter a oportunidade de estudar, novamente, em uma universidade pública, gratuita, inclusiva e de excelência.

Em especial, agradeço à professora Dra. Kelly Bruch, cuja orientação foi fundamental para a conclusão deste trabalho em um momento histórico tão singular. Obrigado pela gentileza, humildade e dedicação, qualidades que fizeram essa jornada se transformar em um aprendizado ainda mais importante para mim.

Aos amigos cariocas e gaúchos, como a Ana, Anna, Bárbara, Ju, Luiz, Rosi e Thamires, pela demonstração de que estar perto não é mais uma grandeza física, o que foi tão importante nos últimos meses.

À minha família, especialmente meus avós, pais, Isabella e Erick pela compreensão diante da minha ausência. Obrigado pelo apoio material e afetivo que possibilitaram a realização deste e de todos os outros projetos.

*“Não se pode esperar encontrar explicações para o que se julga "ruim", para um "mau funcionamento" da sociedade, quando não se é capaz de explicar, ao mesmo tempo, aquilo que se avalia como "bom", "normal" ou "funcionando bem", e vice-versa. (...) Isso também se aplica às crenças distorcidas (...), damos a isso o nome de "preconceito" e o consideramos digno de investigação, talvez na esperança de que seja possível fazer alguma coisa a respeito dele no final. Mas o certo é que não se conseguirá fazer nada sobre ele se o "preconceito" for estudado em isolamento, sem referência a toda a configuração em que ocorre.” Norbert Elias, “Os Estabelecidos e os Outsiders”*

## RESUMO

Os algoritmos formam a estrutura lógica de funcionamento de programas de computador e atualmente compõem sistemas de inteligência artificial (IA) de aplicação ampla na sociedade. Porém, algoritmos enviesados representam sérios riscos de discriminação para seus usuários, que podem produzir e/ou amplificar situações de racismo, sexismo, ou mesmo violações de direitos do consumidor, dentre outras. Por exemplo, existem sistemas de reconhecimento facial que consideram que pessoas não-brancas tem maior chance de reincidência criminal, a partir da análise do tom de pele. Com a disseminação de ferramentas de IA cada vez mais sofisticadas, a sociedade demandará respostas institucionais para os problemas causados pela discriminação algorítmica. Neste cenário, os operadores do Direito precisam identificar o método mais eficaz de concretizar o mandamento presente no Art. 3º da Constituição, que elenca a não-discriminação como um objetivo fundamental do Estado. Considerando que o Brasil não dispõe de uma política pública especificamente voltada para este tema, este trabalho busca identificar maneiras de proporcionar eficácia ao texto constitucional e garantir a não-discriminação no uso da IA. Assim, a pesquisa tem por objetivo entender se e, em caso positivo, como o Direito brasileiro lida com os casos de discriminação causados pelo uso de algoritmos enviesados, especialmente em relação à (desobrigação de) revisão humana de decisões automatizadas. Deste modo, este trabalho analisa o sistema regulatório vigente para identificar eventuais insuficiências normativas. Para atingir este objetivo é adotada metodologia de caráter exploratório-dedutivo, por meio da análise e revisão da literatura científica a respeito do tema, além dos dispositivos institucionais aplicáveis (legislação e jurisprudência). Por fim, a conclusão apresenta a regulação da IA enquanto possível estratégia para oferecer condições que possam promover a eficácia das normas constitucionais e restringir práticas discriminatórias. O argumento final parte do pressuposto de que a revisão humana de decisões automatizadas é talvez a melhor ferramenta disponível neste momento para combater a discriminação algorítmica.

Palavras-chave: LGPD. Aprendizado de máquina. Direito da tecnologia. Direito regulatório. Preconceito.

## ABSTRACT

Algorithms make up the logical structure of computer's software functioning and currently compose artificial intelligence (AI) systems of wide application in society. However, biased algorithms represent serious risks of discrimination for their users, which can produce and / or amplify situations of racism, sexism, or even violations of consumer rights, among others. For instance, there are facial recognition systems that consider that non-white people have a greater chance of criminal recidivism, based on the analysis of skin tone. With the spread of increasingly sophisticated AI tools, society will demand institutional responses to the problems caused by algorithmic discrimination. In this scenario, legal operators will need to identify the most effective method of implementing the mandate contained in Article 3 of the Brazilian Federal Constitution, which lists non-discrimination as a fundamental objective of the State. Considering that Brazil does not have a public policy specifically focused on this topic, this paper seeks to answer the question of how to make the constitutional text effective and ensure non-discrimination in the use of AI. Thus, the research aims to understand whether and, if so, how Brazilian law deals with cases of discrimination caused by the use of biased algorithms, especially in relation to the (legal release from) human review of automated decisions. In this way, this work analyzes the current regulatory system to identify any regulatory insufficiencies. To achieve this objective, an exploratory-deductive methodology is adopted, through the analysis and review of the scientific literature on the subject, in addition to the applicable institutional provisions (legislation and jurisprudence). Finally, the conclusion presents the regulation of AI as a possible strategy to offer conditions that can promote the effectiveness of constitutional norms and restrict discriminatory practices. The final argument is based on the assumption that the human review of automated decisions is perhaps the best tool available at this time to combat algorithmic discrimination.

Keywords: LGPD. Machine learning. Technology law. Regulatory law. Prejudice.

## **LISTA DE ABREVIATURAS E SIGLAS**

API - *Application Programming Interface* (Interface de Programação de Aplicação)

IA – Inteligência Artificial

LGPD – Lei Geral de Proteção de Dados

RGPD – Regulamento Geral de Proteção de Dados (da União Europeia)

## LISTA DE FIGURAS E TABELAS

Figura 1: Exemplo de árvore de decisão.....	29
Figura 2: Proporção de profissionais de tecnologia que presenciaram decisões em seu local de trabalho que poderiam provocar consequências negativas para as pessoas e a sociedade ....	60
Figura 3: Percentual de entrevistados que concordam com a seguinte afirmação: "o Facebook é a internet" .....	60
Tabela 1: Exemplos concretos de racismo algorítmico categorizados pelo conteúdo específico de discriminação racial que apresentam .....	49

## SUMÁRIO

1 INTRODUÇÃO.....	13
2 INTELIGÊNCIA ARTIFICIAL.....	16
2.1 A LÓGICA.....	16
2.2 AS ORIGENS DA CIÊNCIA DA COMPUTAÇÃO.....	17
2.3 O DESENVOLVIMENTO DA INTELIGÊNCIA ARTIFICIAL.....	20
2.4 ALGORITMO E SISTEMAS ALGORITMOS.....	26
3 CONTEXTO DA PESQUISA.....	31
4 VIÉS ALGORÍTMICO E DISCRIMINAÇÃO.....	40
4.1 VIÉS ALGORÍTMICO.....	41
4.2 DISCRIMINAÇÃO ALGORÍTMICA.....	48
5 POSSÍVEIS SOLUÇÕES REGULATÓRIAS PARA O BRASIL.....	56
5.1 A DIMENSÃO ÉTICA.....	57
5.2 REGULAÇÃO JURÍDICA.....	61
6 CONCLUSÃO.....	67
REFERÊNCIAS.....	69

## 1 Introdução

As inovações tecnológicas aceleram a difusão de novas ferramentas que podem apresentar grande utilidade para as pessoas, mas que eventualmente trazem consigo efeitos inesperados e indesejados. O estágio de desenvolvimento tecnológico atual estabelece condições para mudanças sociais profundas, ao oferecer novas formas de interações interpessoais, novos tipos de mobilidade e transporte, ou novas formas de geração de energia, por exemplo. Ainda, pela primeira vez na história a sociedade vem se familiarizando ao convívio com dispositivos (como robôs de limpeza e de atendimento virtual por texto e voz, *drones* e outras formas de inteligência artificial) que não possuem vida, mas se apresentam cada vez mais autônomos, inteligentes e capazes de aprender por si mesmos.

Entretanto, essas novas aplicações da tecnologia moderna em geral vêm acompanhadas de incertezas no que diz respeito aos seus efeitos de longo prazo. Na medida em que ocorre o desenvolvimento de novas aplicações tecnológicas, também se verificam mudanças na percepção da sociedade a respeito do seu uso e da correspondente resposta regulatória que lhe seria adequada. Diferentes sociedades e seus respectivos cidadãos possuem diversas preferências coletivas e/ou individuais a respeito do grau de incerteza e do tipo de risco que são considerados aceitáveis. Neste contexto, o avanço tecnológico oferece a oportunidade de realizar algumas reflexões relevantes, como, por exemplo, a que diz respeito ao melhor alinhamento possível entre a tecnologia e os valores sociais e humanos considerados universais. Também se pode questionar se neste momento a humanidade realmente tem consciência acerca dos valores mais relevantes que devem ser promovidos (e se há consenso mínimo a respeito desses valores), ou ainda como garantir que o mundo digital se torne melhor do que o mundo analógico em relação aos indivíduos, à sociedade e ao planeta. A maneira pela qual a regulação normativa promovida pelo Estado pode atender a esses objetivos sociais diversos e por vezes conflitantes permanece uma questão fundamental da ciência jurídica. Novas tecnologias também levantam questionamentos a respeito dos limites do Direito na medida em que se torna difícil identificar e separar o que é prejudicial do que é benéfico. Também é preciso considerar que a aceitação social das novas tecnologias é essencial para que elas sejam bem-sucedidas.

Como o título sugere, este trabalho trata das respostas regulatórias (brasileiras) para a questão da discriminação algorítmica. Neste sentido, o problema de pesquisa aqui é: a ausência de revisão humana das operações automatizadas é prejudicial para se mitigar a discriminação algorítmica? Ou seja, diante da conjuntura apresentada nos parágrafos anteriores, é preciso analisar a capacidade de a discriminação algorítmica no Brasil ser corrigida sem que exista

obrigatoriedade de intervenção humana nas decisões automatizadas. Por conseguinte, o objetivo geral desta pesquisa consiste em identificar maneiras de conferir maior eficácia ao texto constitucional brasileiro e garantir a não-discriminação no uso da inteligência artificial. Isto é, este trabalho busca estudar se a atual desobrigação legal de revisão humana sobre as decisões automatizadas compromete a eficácia das disposições constitucionais e a proteção aos direitos humanos no Brasil nos casos que envolvem especialmente a discriminação algorítmica em âmbito nacional. Conforme indicado no terceiro capítulo deste trabalho, no caso da União Europeia (modelo de inspiração para o Brasil nesta área) foi expressamente estabelecida a revisão, por operadores humanos, de decisões automatizadas efetuadas por algoritmos. Uma análise mais profunda poderia avaliar mesmo a capacidade de toda a legislação aplicável ao tema para coibir violações de direitos, em especial direitos fundamentais, causadas por operações automatizadas realizadas por algoritmos enviesados.

A justificativa para esta pesquisa está no fato de que o uso da ciência de dados já atinge patamares que seriam considerados como literatura de ficção há alguns anos.<sup>1</sup> Neste cenário, é imprescindível realizar uma reflexão sobre o uso dos dados que alimentam os algoritmos de inteligência artificial, sobretudo quando se acredita, por vezes, na suposta infalibilidade da tecnologia. Esse assunto se coloca como um tema de estudos tanto interessante quanto desafiador dentro do Direito, pela dificuldade e atualidade do tema. A Universidade Federal do Rio Grande do Sul, por ser um dos centros de ensino e pesquisa de grande comprometimento com a inovação no cenário acadêmico brasileiro, é uma instituição que pode propiciar as condições para o desenvolvimento de estudos sobre este tema, devido à diversidade e capacidade de sua comunidade acadêmica.

Por fim, este trabalho tem a intenção de ir além de uma simples reflexão abstrata, procurando analisar fenômenos concretos e contribuir para o desenho de soluções eficazes, em especial considerando a atuação do poder público. Por se tratar de um tema novo para a academia brasileira, não se espera identificar vasta bibliografia específica sobre o objeto desta pesquisa em língua portuguesa, de modo que as fontes consultadas serão em grande medida artigos acadêmicos estrangeiros (de diversas áreas do conhecimento), além da doutrina jurídica brasileira. Assim, é adotado o método dedutivo, com caráter descritivo-exploratório, mediante

---

<sup>1</sup> Como por exemplo, o uso de aplicativos em países como China e Coreia do Sul que monitoram os cidadãos e informam a necessidade de quarentena ou a autorização para circular pela cidade durante o surto de COVID-19 nos primeiros meses de 2020, conforme demonstram os links a seguir: <https://www.nytimes.com/2020/03/01/business/china-coronavirus-surveillance.html>; <https://coronamap.site/>. Acesso em 16 de março de 2020.

análise da bibliografia científica publicada sobre o tema para identificar a interseção entre os artigos acadêmicos das áreas de direito e ciência da computação, sobretudo.

Entretanto, uma particularidade do desenvolvimento deste trabalho diz respeito ao caráter inicialmente descritivo que o texto apresenta, em especial nesta introdução e nos dois capítulos seguintes, que apresentam alguns conceitos importantes relacionados à área da inteligência artificial e a contextualização do tema. Ao longo dos capítulos subsequentes, a redação assumirá perspectiva cada vez mais exploratória, na medida em que se busca compreender o fenômeno da discriminação algorítmica e seu tratamento jurídico atual. Isso se deve em grande parte ao fato de que o objeto dessa pesquisa corresponde a uma situação relativamente recente, que há pouco tempo vem recebendo atenção significativa por parte da academia. Ainda, dada a atualidade deste assunto, também serão analisadas notícias e eventuais discussões legislativas que abordem este tópico.

Com a finalidade de atender aos objetivos desta pesquisa, o texto foi organizado em diferentes capítulos que tratam dos objetivos específicos da pesquisa: esta introdução, que inclui a apresentação do problema a ser enfrentado, os objetivos gerais e específicos, além da metodologia da pesquisa; a exposição sobre o desenvolvimento e funcionamento dos mecanismos mais difundidos de inteligência artificial (capítulo 2); a contextualização da pesquisa (capítulo 3) a análise dos principais tipos de vieses que originam situações discriminatórias em mecanismos de inteligência artificial, além de tratar do próprio conceito de discriminação (capítulo 4). Em seguida, é realizada avaliação a respeito da aptidão dos mecanismos jurídicos disponíveis no Brasil para lidar com os problemas associados ao viés algorítmico (capítulo 5). Por fim, a conclusão da pesquisa traz o argumento principal de que a revisão humana de decisões automatizadas talvez seja melhor ferramenta disponível neste momento para combater a discriminação algorítmica.

## 2 Inteligência artificial

O tema deste trabalho envolve alguns conceitos principais (como inteligência artificial e algoritmo) que devem ser previamente abordados para que se possa tratar de modo mais detalhado do viés algorítmico e suas implicações jurídicas. É importante observar, entretanto, que os termos apresentados aqui privilegiam a perspectiva jurídica, sem grande aprofundamento técnico relativo à área da ciência da computação. Conforme indicado anteriormente, a intenção é apenas facilitar a compreensão do viés algorítmico por profissionais e estudantes de direito, para em um segundo momento analisar seus aspectos jurídicos. Deste modo, o objetivo específico deste capítulo é apresentar o desenvolvimento e funcionamento dos principais mecanismos de inteligência artificial, o que inclui por exemplo seu desenvolvimento histórico e sua base de funcionamento, o algoritmo.

Para atingir esse objetivo, é oportuno conhecer de forma sucinta a história da computação, disciplina científica que alcançou sua maturidade no século passado. O intuito é ir além de mera enumeração de fatos ou do crédito a determinados pesquisadores relevantes desta área. Trata-se de descobrir e traçar o modo de pensar que levou a importantes descobertas científicas que hoje são fundamentais para a sociedade. Neste aspecto, torna-se interessante lembrar autores e publicações em que, pela primeira vez, foram introduzidos alguns conceitos, teorias, métodos e ferramentas que provocaram grandes mudanças tecnológicas.

Deste modo, essa seção traz algumas noções sobre a lógica, para em seguida apresentar breve histórico da ciência da computação. Após, o texto aborda o desenvolvimento da inteligência artificial e por fim os algoritmos e seus sistemas.

### 2.1 A lógica

Os conceitos abordados adiante têm por fundamento a lógica da computação, que se desenvolveu imensamente durante os últimos cem anos, a partir de teorias matemáticas apresentadas por diversos estudiosos europeus desde o século XVII.<sup>2</sup> Em geral, a lógica é um

---

<sup>2</sup> A Lógica, tal qual “descoberta” por Aristóteles na Grécia Antiga, sofreu profundas alterações a partir dos estudos de Gottfried Wilhelm Leibniz (1646-1716), Johann Heinrich Lambert (1728 – 1777) e, principalmente, George Boole (1815 – 1864), Augustus De Morgan (1806 – 1871), Alfred North Whitehead (1861 – 1947) e Bertrand Russell (1872 – 1970), no que tange à suas aplicações na área da computação. Entretanto, nas últimas décadas foi demonstrado que a Matemática não é uma ciência una, possuindo diferentes possibilidades teóricas que podem ser consideradas em qualquer nova aplicação prática. ABE, Jair Minoro; SCALZITTI, Alexandre; SILVA FILHO, João Inácio. *Introdução à lógica para a ciência da computação*. São Paulo: Arte e Ciência, 2002, pp. 11-12.

ramo da filosofia que possui difícil definição. Grosso modo, seu marco inicial é associado ao pensamento de Aristóteles na antiguidade clássica, que buscou organizar a forma de pensar dos seres humanos. As regras da lógica determinariam, por conseguinte, a melhor maneira de se atingir a conclusão adequada a partir das premissas disponíveis, para cada ciência e seus respectivos objetos de estudo. Trata-se, portanto, de uma disciplina propedêutica<sup>3</sup>, isto é, que busca compreender e reger corretamente o pensamento humano (para prepará-lo para analisar adequadamente um problema específico). Na Idade Moderna, a lógica exerceria papel fundamental no estabelecimento do método científico (que utiliza evidências empíricas interpretadas pela razão humana para construir o conhecimento).

No que tange à ciência da computação, a lógica atualmente possui algumas ramificações que foram especialmente relevantes para o desenvolvimento de computadores e seus sistemas de operação. Como exemplo, é possível citar a teoria da recursão, que trata do que é exequível mecanicamente, sem intervenção do intelecto humano. A máquina desenvolvida por Alan Turing durante a Segunda Guerra Mundial (que será abordada adiante) é uma realização física desta teoria. Há ainda outros ramos, como aquele que analisa os fundamentos da própria Matemática, além da lógica algébrica, que contribuíram bastante para o desenvolvimento de aplicações tecnológicas e que não serão abordados aqui por extrapolarem o escopo principal deste trabalho.<sup>4</sup>

## 2.2 As origens da ciência da computação

Em seu sentido original, no início do século passado o termo “computador” designava os profissionais que trabalhavam realizando cálculos segundo métodos eficazes, tarefa que hoje é desempenhada pelos atuais computadores eletrônicos. Esses profissionais trabalhavam em diversos estabelecimentos, como órgãos públicos, empresas e centros de pesquisa. Foi só a partir da década de 1920 que se tornou comum o uso da expressão “máquina de computação” (*computing machine*) para fazer referência aos aparelhos capazes de executar o trabalho de um computador (humano). Após a Segunda Guerra Mundial, com o aparecimento das máquinas de computação eletrônica, progressivamente se adotou a denominação “computador” de modo

---

<sup>3</sup> ARANHA, Maria Lúcia de Arruda; MARTINS, Maria Helena Pires. *Filosofando : Introdução à Filosofia*. São Paulo : Moderna, 1986. p. 96.

<sup>4</sup> ABE, Jair Minor; SCALZITTI, Alexandre; SILVA FILHO, João Inácio. *Idem*, ibidem. pp. 15-16.

genérico para esses instrumentos, que no início ainda eram acompanhados do termo “eletrônico” ou “digital”.<sup>5</sup>

Mas os computadores de hoje são herdeiros de uma longa tradição científica que se intensificou durante a busca pela automação de atividades e melhoria de processos industriais que caracterizaram a Europa ocidental do século XIX. É nessa conjuntura histórica que o matemático inglês Charles Babbage (1791-1871) se tornou talvez o principal precursor da história da computação, demonstrando grande paixão por máquinas de cálculo durante seu trabalho acadêmico. Sua principal invenção, chamada de “máquina diferencial”, consistiu na primeira tentativa bem-sucedida de construir uma calculadora automática (e que permanece como um dos melhores exemplos de engenharia de precisão de seu tempo), capaz de compilar tabelas matemáticas de trigonometria e logaritmos. Posteriormente, Babbage projetou um instrumento ainda melhor que seria capaz de executar quaisquer tipos de cálculos, a “máquina analítica”, que já apresentava algumas das características dos computadores de hoje, conforme descrição<sup>6</sup> a seguir:

A Máquina Analítica proposta por Babbage, consideravelmente mais ambiciosa que a Máquina Diferencial, teria sido um computador digital mecânico de uso geral. A Máquina Analítica teria um armazenamento de memória e uma unidade central de processamento (ou 'moinho') e poderia selecionar dentre as ações alternativas conforme o resultado de suas ações anteriores (uma instalação hoje conhecida como ramificação condicional). O comportamento da Máquina Analítica seria controlado por um programa de instruções contidas em cartões perfurados conectados com fitas (uma ideia que Babbage adotara do tear de tecelagem de Jacquard). Babbage enfatizou a generalidade da Máquina Analítica, dizendo que "as condições que possibilitam a uma máquina finita fazer cálculos de extensão ilimitada estão presentes na Máquina Analítica." (tradução livre)

---

<sup>5</sup> COPELAND, B. Jack, *The Modern History of Computing*. The Stanford Encyclopedia of Philosophy (Winter 2017 Edition), Edward N. Zalta (ed.). Disponível em: <https://plato.stanford.edu/entries/computing-history/>. Acesso em 3 de julho de 2020.

<sup>6</sup> “Babbage's proposed Analytical Engine, considerably more ambitious than the Difference Engine, was to have been a general-purpose mechanical digital computer. The Analytical Engine was to have had a memory store and a central processing unit (or ‘mill’) and would have been able to select from among alternative actions consequent upon the outcome of its previous actions (a facility nowadays known as conditional branching). The behaviour of the Analytical Engine would have been controlled by a program of instructions contained on punched cards connected together with ribbons (an idea that Babbage had adopted from the Jacquard weaving loom). Babbage emphasised the generality of the Analytical Engine, saying ‘the conditions which enable a finite machine to make calculations of unlimited extent are fulfilled in the Analytical Engine’. COPELAND, B. Jack. *Idem*, *ibidem*.”

No entanto, devido a dificuldades orçamentárias e mesmo tecnológicas da época (já que os protótipos de Babbage eram muito avançados para o estágio de desenvolvimento científico de meados do século XIX), os experimentos não foram totalmente concluídos na prática.<sup>7</sup>

É importante lembrar que o alcance obtido pelo segundo invento de Charles Babbage não seria possível sem a contribuição visionária de Augusta Ada Byron King, Condessa de Lovelace (1815-1852, que era filha do poeta inglês Lord Byron, sendo conhecida atualmente como Ada Lovelace), considerada a primeira mulher programadora da história. Babbage trabalhou em estreita colaboração com Lovelace, em cuja homenagem foi nomeada a atual linguagem de programação ADA. Ada Lovelace tinha uma abordagem peculiar em relação à matemática, tratando esta ciência como uma espécie de poesia. Essa inusitada perspectiva de caráter artístico permitiu que a cientista elaborasse suas previsões a respeito da possibilidade de usar a Máquina Analítica de Babbage para computação não numérica, sugerindo que o mecanismo poderia até ser capaz de compor peças elaboradas de música<sup>8</sup> (o que de fato se concretizaria no futuro).<sup>9</sup>

Durante a primeira metade do século XX, as máquinas que precederam o atual computador seguiam o modelo mecânico e analógico da Máquina Analítica de Babbage, ou seja, operavam basicamente enquanto representações físicas de funções matemáticas que relacionavam duas grandezas. Por meio de uma analogia realizada pelo usuário (treinado), era possível compreender o resultado da atividade da máquina.<sup>10</sup> Mas o grande impulso se deu com o desenvolvimento do modelo digital de computação.<sup>11</sup>

---

<sup>7</sup> Entretanto, a Máquina Analítica chegou a ser concluída pelos suecos Georg e Edvard Scheutz em 1854, ainda durante a vida do inventor inglês. Essas e outras informações a respeito da vida e obra de Charles Babbage estão disponíveis na página eletrônica do instituto que leva seu nome (vinculado à Universidade de Minnesota), nos Estados Unidos: <http://www.cbi.umn.edu/about/babbage.html>. Acesso em 3 de julho de 2020.

<sup>8</sup> COPELAND, B. Jack. Idem, ibidem.

<sup>9</sup> Como exemplo é possível citar o projeto Iamus, desenvolvido pela Universidade de Málaga, na Espanha. Detalhes disponíveis em: <http://www.geb.uma.es/melomics/melomics.html>. Acesso em 4 de julho de 2020.

<sup>10</sup> Um exemplo simples de máquina analógica é o relógio analógico, que recebe seu nome por dividir o tempo da rotação terrestre de modo a possibilitar sua mensuração específica pelos indivíduos. A movimentação dos ponteiros do relógio analógico também evidencia o fato de que as máquinas analógicas apresentam valores de modo contínuo e proporcional às variações da grandeza física mensurada (o que é matematicamente representado por uma função linear). No caso do relógio, cada volta completa do ponteiro menor representa 12 horas (meio dia ou meia rotação terrestre). O mesmo raciocínio se aplica para o ponteiro maior, que se desloca proporcionalmente (12 vezes) mais rápido de modo que cada volta completa representa uma hora (ou 1/24 de um dia ou rotação terrestre).

<sup>11</sup> Existem vários cientistas que atuaram (por vezes de modo independente) a partir da década de 1930 para criar um modelo digital de computação. Aqui vale ressaltar nomes como Claude Shannon (1916-2001) e Akira Nakashima (1908-1970), cujos trabalhos foram essenciais para o desenvolvimento do *hardware* (parte física do computador) que tornaria viável o aumento da capacidade de processamento dos computadores e consequentemente, a migração para o modelo digital na computação, fundamentando grande parte dos avanços tecnológicos da segunda metade do século XX. Shannon demonstrou que uma aplicação elétrica utilizando álgebra booleana poderia resolver qualquer problema de lógica. Sua tese de mestrado está disponível (em inglês) no seguinte endereço eletrônico: <https://www.cs.virginia.edu/~evans/greatworks/shannon38.pdf>. Acesso em 4 de

### 2.3 O desenvolvimento da inteligência artificial

As investigações científicas que deram origem à ciência da computação tal como a conhecemos hoje se iniciaram no contexto histórico que produziu a Segunda Guerra Mundial. Durante o conflito, a tentativa do governo britânico de decifrar a criptografia alemã no intuito de se antecipar aos ataques aéreos e navais inimigos se tornou o grande fomento às pesquisas de Alan Turing, matemático inglês contratado pelos militares de seu país para auxiliar no esforço de guerra. A máquina desenvolvida por Turing foi capaz de cumprir esse objetivo, e seu trabalho hoje é reconhecido por ter abreviado o conflito e contribuído diretamente para evitar milhões de mortes no continente europeu.

Após o fim da guerra, Alan Turing publicou sua descoberta em um artigo científico<sup>12</sup> que seria considerado o marco histórico da invenção do atual computador. Naquele momento Turing também conseguiu antever a capacidade extraordinária de sua máquina<sup>13</sup>, bem como o desenvolvimento rápido que ela apresentaria nos anos seguintes. Deste modo, o cientista imaginou um teste (que futuramente levaria seu nome<sup>14</sup>) que fosse capaz de discernir e separar

---

julho de 2020; Já Nakashima estabeleceu as bases do circuito eletrônico integrado. Para detalhes a respeito de sua obra, consultar: Radomir S. Stanković, Jaakko Astola (2008), *Reprints from the Early Days of Information Sciences: TICSP Series On the Contributions of Akira Nakashima to Switching Theory*, TICSP Series #40, Tampere International Center for Signal Processing, Tampere University of Technology. Disponível (em inglês) em: <http://ticsp.cs.tut.fi/reports/reprint-nakashima-rr.pdf>. Acesso em 4 de julho de 2020.

<sup>12</sup> TURING, Allan M. Computing Machinery and Intelligence. *Mind*, v. 59, n. 236, 1950, pp. 433-460. Disponível em: <https://academic.oup.com/mind/article-abstract/LIX/236/433/986238>. Acesso em 29 de fevereiro de 2020.

<sup>13</sup> Em entrevista ao jornal espanhol El País, o professor da Faculdade de Informática da Universidade Complutense de Madri, Ricardo Peña Marí, esclareceu que “a Máquina de Turing não é uma máquina que existe no mundo físico, mas um constructo mental. Consiste em uma faixa infinita sobre a qual vão sendo feitas operações repetitivas até oferecer soluções, vem a ser uma definição informática do algoritmo e de um computador, o primeiro, conceitualizado: “Em essência, é o precursor dos computadores: tem memória, instruções (um programa), operações elementares, uma entrada e uma saída”, explica o professor Peña. O mais interessante é que é uma máquina universal, que pode executar qualquer programa que lhe for solicitado. Dentro dos problemas do mundo existem dois tipos: aqueles que podem ser resolvidos por uma Máquina de Turing (chamados computáveis) e aqueles que não podem (os não computáveis), assim como vemos no mundo real tarefas que as máquinas podem executar (cada vez mais) e outras que só os humanos podem realizar. Todos os computadores, tablets, smartphones, etc., que conhecemos são máquinas de Turing.” Disponível em: [https://brasil.elpais.com/brasil/2018/03/30/tecnologia/1522424604\\_741609.html](https://brasil.elpais.com/brasil/2018/03/30/tecnologia/1522424604_741609.html). Acesso em 7 de junho de 2020.

<sup>14</sup> “Turing suggested we postpone a direct answer to the question whether machines can think; he proposed that we ask instead whether an artifact could fool a series of questioners as often as the human was able to convince them of the truth, about half the time. The advantage of Turing's test is that it avoids direct confrontation with the difficult questions about what "thinking" or "intelligence" is. Turing thought that he had devised a test that was so difficult that anything that could pass the test would necessarily qualify as intelligent” (LAWRENCE, B. Solum. *Legal Personhood for Artificial Intelligences*. North Carolina Law Review, v. 70, n. 4, 1992. p. 1236). Disponível em: <https://scholarship.law.unc.edu/cgi/viewcontent.cgi?article=3447&context=nclr>. Acesso em 16 de março de 2020.

a inteligência humana da inteligência do computador (inteligência artificial), isto é, um meio para identificar e diferenciar pessoas de máquinas<sup>15</sup>.

A principal prerrogativa do teste de Turing consistiu em evitar abordar diretamente a natureza da inteligência humana e a possibilidade de sua reprodução em máquinas; o cientista preferiu concentrar seus esforços em mensurar a habilidade do computador de enganar avaliadores e convencer humanos de sua inteligência, sem investigar a real natureza de seu processamento de informações ou, em outras palavras, a capacidade do computador de “pensar” (e suas habilidades cognitivas de modo geral). O teste era relativamente simples, ao qual Turing deu o nome de “jogo da imitação”: a máquina seria avaliada junto a um oponente humano, por um avaliador que fazia perguntas de qualquer natureza a ambos (humano e máquina), sem saber qual é qual. Aqui, tanto o humano quanto a máquina se passariam por humanos em suas respostas, tentando convencer o avaliador de que o outro jogador seria a máquina. Após o fim da rodada de perguntas, o avaliador deveria decidir qual jogador seria o humano. Turing acreditava que seu “jogo da imitação” era tão complexo que todo ente que fosse aprovado (isto é, que conseguisse convencer o avaliador de que seria humano) necessariamente seria classificado como um ser inteligente.

Deste modo, o teste de Turing representou a primeira tentativa na definição do conceito de inteligência artificial, em um momento em que a própria ciência da computação estava em sua fase inicial. Ao longo da segunda metade do século XX, as pesquisas científicas levaram a uma multiplicidade de definições para este conceito<sup>16</sup>, já que diversos estudiosos enfatizaram perspectivas diferentes a respeito deste tema. Além disso, na medida em que o desenvolvimento tecnológico produz computadores cada vez mais sofisticados, as atividades que já foram consideradas como exemplos de inteligência artificial agora são encaradas como triviais e excluídas da compreensão contemporânea a respeito do que seria uma “máquina inteligente”.<sup>17</sup>

---

<sup>15</sup> Um modelo do teste de Turing é frequentemente aplicado em sites de tribunais para consulta de processos, mas de modo reverso (o computador aplica o teste para o usuário humano), por exemplo, onde o usuário deve digitar os números que aparecem em uma imagem distorcida, chamada de *captcha*. Entretanto, já se sabe que computadores mais modernos conseguem passar neste modelo de teste, como demonstra o artigo a seguir: YE, Guixin; TANG, Zhanyong; FANG, Dingyi; ZHU, Zhanxing; FENG, Yansong; XU, Pengfei; CHEN, Xiaojiang; WANG, Zheng. *Yet Another Text Captcha Solver: A Generative Adversarial Network Based Approach*. In: Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security (CCS '18). New York: Association for Computing Machinery, 2018. p. 332–348. Disponível em: <https://dl.acm.org/doi/10.1145/3243734.3243754>. Acesso em 16 de março de 2020.

<sup>16</sup> ROTHKEGEL, Tobias; TAYLOR, Mark. *What characterizes artificial intelligence and how does it work?* Computer and Telecommunications Law Review, v. 22, n. 4, 2016. p. 98. (apud MACHADO, E. S. P. 2019).

<sup>17</sup> Esse fenômeno é conhecido pelo termo “efeito IA” (de “inteligência artificial”), onde os avanços no desenvolvimento da inteligência artificial passam a ser considerados como inseridos dentro da capacidade genérica de processamento do computador. Para mais detalhes, consultar: MCCORDUCK, Pamela. *Machines who think: a personal inquiry into the history and prospects of artificial intelligence*. AK Peters Ltd: Natick, 2004. p. 423.

Entretanto, considerando tanto a abrangência quanto os objetivos deste trabalho, essa discussão não será aprofundada. Aqui, a inteligência artificial pode ser definida genericamente como um sistema (baseado em algoritmos, o que também será analisado adiante) que realiza atividades complexas que normalmente requerem a cognição humana para sua consecução, tal como o aprendizado, a comunicação e a tomada de decisões. Assim, as máquinas dotadas de inteligência artificial são programadas (inicialmente por seres humanos) para aprender por si mesmas, melhorando de modo autônomo seu desempenho em atividades normalmente executadas por pessoas. O chamado aprendizado de máquina (também muito conhecido por seu nome em inglês, *machine learning*) se refere justamente a essa característica da inteligência artificial, que vem apresentando inúmeras aplicações tecnológicas nas últimas décadas.

Seguindo o rigor técnico da ciência da computação, o que se entende comumente por “aprendizado de máquina” na verdade abrange uma série de sistemas bastante diferentes entre si. Entretanto, neste texto essas diferenças técnicas não serão abordadas, pois neste momento é possível realizar uma generalização sem prejuízos para a compreensão deste trabalho. Aqui, o aprendizado de máquina é tratado genericamente, e corresponde provavelmente ao mecanismo de inteligência artificial mais difundido atualmente.<sup>18</sup> Seu propósito é permitir a um sistema de algoritmos aprender a partir dos subsídios fornecidos pela análise de dados (em massa, sobretudo), sem haver programação específica para cada aprendizado em si.<sup>19</sup> Em outras palavras, é uma ferramenta capaz de permitir o aprendizado do computador sem a intervenção humana. Grosso modo, pode-se dizer que esse aprendizado ocorre basicamente a partir de uma análise estatística realizada sobre o conjunto de dados disponível, permitindo ao sistema algorítmico interagir com os dados (de modo constante, mesmo que por meio de conexão de rede ou Internet), realizando um tratamento e descrição das informações, além de identificar padrões no conjunto de dados e com isso prever os resultados de suas interações.

Segundo a definição técnica da Organização Mundial da Propriedade Intelectual, o aprendizado de máquina é um processo de inteligência artificial que usa algoritmos e modelos estatísticos para permitir que computadores tomem decisões sem precisar que sejam programados explicitamente para executar a tarefa. Os algoritmos do aprendizado de máquina constroem um modelo em dados de amostragem usados como dados de treinamento para identificar e extrair padrões da universalidade de dados disponíveis e, portanto, adquirir seu

---

<sup>18</sup> SURDEN, Harry. *Machine Learning and Law*. Washington Law Review, v. 89, 2014. p. 89. Disponível em: <https://digitalcommons.law.uw.edu/wlr/vol89/iss1/5/> Acesso em 16 de março de 2020.

<sup>19</sup> HURWITZ, Judith; KIRSCH, Daniel. *Machine Learning For Dummies*: IBM Limited Edition. Hoboken: John Wiley & Sons Inc., 2018. Disponível em: <https://www.ibm.com/downloads/cas/GB8ZMQZ3>. Acesso em 16 de março de 2020.

próprio conhecimento.<sup>20</sup> Neste sentido, afirma-se que o aprendizado ocorreu no momento em que o sistema obtém o resultado gerado a partir das interações do algoritmo com o conjunto de dados inseridos. Como em qualquer análise estatística, o aumento progressivo na quantidade e na qualidade de informações a respeito desse conjunto de dados acarretará um resultado mais fidedigno em relação à realidade, isto é, a precisão do sistema aumentará na medida em que mais dados forem disponibilizados. Em suma, vastas quantidades de dados, desde que precisas e organizadas (em geral por programadores humanos), podem ser alimentadas ao sistema algorítmico para que este aprenda com as informações inseridas e consiga gerar um modelo de atuação que poderá ser replicado futuramente em outros bancos de dados.

Assim, o aprendizado de máquina pode ser entendido como a habilidade do sistema de aprender e adaptar a sua própria programação inicial, obtendo resultados por vezes imprevisíveis até mesmo para seus programadores humanos, ou seja:<sup>21</sup>

“Os programas de aprendizado de máquina têm a capacidade de alterar ou adaptar sua própria programação com base em novos dados, à medida em que esses dados são apresentados ao programa ou descobertos pelo próprio programa (por meio de pesquisas na web, entrada de câmera etc.). O aprendizado de máquina representa a possibilidade de um programa de IA que cria algo totalmente inesperado e não intencional pelo programador original, e que pode não precisar de nenhum envolvimento humano adicional para ser considerado um trabalho comercialmente valioso.” (tradução livre).

Além do aprendizado de máquina,<sup>22</sup> existem outros sistemas sofisticados de inteligência artificial que são de uso comum atualmente. As redes neurais compreendem mais uma técnica que faz uso de um amplo conjunto de mecanismos e modelos de aprendizado de máquina, cuja denominação se deve à sua capacidade de simular o funcionamento do neocórtex humano (a

---

<sup>20</sup> Tradução livre do original: “Machine learning: an AI process that uses algorithms and statistical models to allow computers to make decisions without having to explicitly program it to perform the task. Machine learning algorithms build a model on sample data used as training data in order to identify and extract patterns from data, and therefore acquire their own knowledge. Organização Mundial de Propriedade Intelectual. *WIPO Technology Trends 2019 – Artificial Intelligence*. p. 146. Disponível em: [https://www.wipo.int/edocs/pubdocs/en/wipo\\_pub\\_1055.pdf](https://www.wipo.int/edocs/pubdocs/en/wipo_pub_1055.pdf). Acesso em: 05 de junho de 2020.

<sup>21</sup> “Machine learning programs have the ability to change or adapt their own programming based on new data as it is presented to the program, or discovered by the program itself (through web searches, camera input, etc.). Machine learning represents the possibility of an AI program that creates something totally unexpected and unintended by the original programmer, and that may not need any further human involvement to be considered a commercially valuable work.” WAGNER, James. *Rise of the Artificial Intelligence Author*. Vancouver: The Advocate, v. 75, n. 4, 2017. p. 529. (apud MACHADO, E. S. P. 2019).

<sup>22</sup> Existe uma área emergente de pesquisa em aprendizado de máquina onde o algoritmo não necessariamente interpreta como falso um dado novo, com o qual nunca teve contato. Esse aspecto se torna bastante interessante para que a inteligência artificial consiga de fato lidar com a complexidade do mundo “real” e reduzir alguns tipos de vieses como os que serão apresentados no capítulo 4. Sobre esse assunto, ver: <https://medium.com/@danibrauner/a-sem%C3%A2ntica-de-dados-%C3%A9-um-futuro-poss%C3%ADvel-para-o-aprendizado-de-m%C3%A1quina-9344661b5db7>. Acesso em 31 de outubro de 2020.

parte mais desenvolvida do cérebro). Usando a terminologia técnica da estatística, as redes neurais tem como método principal<sup>23</sup> “extrair combinações lineares dos dados iniciais como características derivadas e depois modelar o alvo como uma função não linear dessas características”, resultando em um “método de aprendizado poderoso, com aplicações difundidas em muitos campos.”

Portanto, as redes neurais seguem os mesmos modelos estatísticos utilizados no aprendizado de máquina, como por exemplo a regressão linear<sup>24</sup>, com a peculiaridade de que absorvem e distribuem sua capacidade de processamento de dados em vários níveis interconectados de processamento, imitando o mecanismo característico dos neurônios no sistema nervoso humano, o que permite identificar conexões e padrões entre as informações contidas nos dados sendo tratados, porém em uma escala bastante ampliada.

As redes neurais, desta forma, apresentam camadas distintas de tratamento de dados para viabilizar o aprendizado de máquina de modo repetitivo (“iterativo”, no jargão da ciência da computação). Os dados são inseridos em uma camada, enquanto sua análise e operação são realizadas por camadas intermediárias, e o sistema disponibiliza os resultados em uma camada final, de saída. O grau de complexidade e sofisticação do sistema acompanhará o número de níveis de processamento que ele possui, isto é, quanto mais níveis, mais sofisticadas serão as redes neurais.

Os sistemas mais complexos de redes neurais são muito conhecidos por seu nome na língua inglesa de *deep learning* (aprendizado profundo), e apresentam cada vez mais possibilidades de aplicação prática, sobretudo em mecanismos de reconhecimento facial e de voz.<sup>25</sup> Esses sistemas são excelentes em reconhecer padrões em dados grandes e não estruturados, como imagens digitais, som ou texto, e em analisar esses padrões para fazer

---

<sup>23</sup> “The central idea is to extract linear combinations of the inputs as derived features, and then model the target as a nonlinear function of these features. The result is a powerful learning method, with widespread applications in many fields.” HASTIE, Trevor; TIBSHIRANI, Robert; FRIEDMAN, J. H. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2ª ed. Nova York: Springer, 2009. p. 389. Disponível em: <https://link.springer.com/book/10.1007%2F978-0-387-84858-7#authorsandaffiliationsbook>. Acesso em 6 de junho de 2020.

<sup>24</sup> Grosso modo, em Estatística se entende o conceito de regressão linear como o(s) modelo(s) de análise matemática para quantificar e compreender a natureza da relação entre magnitudes de duas ou mais variáveis independentes (por exemplo, se X aumenta, Y aumenta, ou ainda se X aumenta, Y diminui). O conceito se aproxima da ideia de correlação, onde o objetivo é mensurar a força do vínculo existente entre duas ou mais variáveis. Essas duas técnicas vêm sendo utilizadas desde o início do século passado como ferramentas de previsão para os resultados ou fenômenos estudados pelos estatísticos. No que tange à ciência de dados, a regressão linear tem exercido papel fundamental no tratamento dos dados, servindo como modelos de previsão e permitindo também a identificação de “anomalias” nos dados que integram a análise. Para maiores detalhes consultar: BRUCE, Andrew; BRUCE, Peter. *Practical Statistics for Data Scientists*. Sebastopol: O’Reilly, 2017. pp. 231-232.

<sup>25</sup> HURWITZ, Judith; KIRSCH, Daniel. Idem, *ibidem*. p. 17. Disponível em: <https://www.ibm.com/downloads/cas/GB8ZMQZ3>. Acesso em 16 de março de 2020.

previsões, sobretudo em situações onde há um volume gigantesco de dados (como interações de milhões de pessoas em redes sociais) ou mesmo quando as informações são muito abstratas ou desorganizadas.

Nesse contexto, os mecanismos de aprendizado profundo conseguem superar os seres humanos em tarefas visuais, como classificação de imagens, reconhecimento facial ou diagnóstico de câncer de pele. Seu desempenho superior representa valiosas oportunidades para que seja possível identificar as ligações entre características dentro dos dados analisados que poderiam ser perdidas ou mal interpretadas pelo cérebro humano, embora seu uso mal-intencionado também signifique riscos na mesma proporção.<sup>26</sup>

Há um outro mecanismo de inteligência artificial que se tornou bastante conhecido nos últimos anos: os sistemas de processamento de linguagem natural<sup>27</sup>, que conferem a capacidade aos computadores de realizar tarefas como o reconhecimento e transcrição da fala humana, além da compreensão da escrita e geração automática de textos. Seu desenvolvimento acompanhou a ciência da computação desde a década de 1950, com avanços consideráveis a partir dos anos 1990. Devido à disseminação desse tipo de tecnologia, hoje é comum que aparelhos de telefonia celular recebam comandos de voz para executar tarefas cada vez mais complexas, por exemplo, ou ainda que o atendimento de clientes pela Internet nos mais diversos estabelecimentos comerciais seja realizado por robôs.

O funcionamento dessa modalidade segue o mesmo modelo das outras formas de inteligência artificial, vistas acima. Os algoritmos trabalham a partir de metodologias que permitem aos programadores mensurar a precisão dos resultados, isto é, se as palavras utilizadas fazem sentido e correspondem à mensagem designada. Deste modo é possível avaliar o código de programação e implementar avanços continuamente. Aqui também se aplicam algoritmos de aprendizado de máquina, para que a ferramenta seja capaz de realizar suas próprias melhorias. As palavras (escritas ou faladas que compõem o resultado) envolvem algum banco de dados, dentro do sistema como um todo. Grosso modo, o processamento de linguagem natural em si pode ser composto por apenas algumas ou mesmo todas as etapas a seguir, dependendo da

---

<sup>26</sup> Já existem estudos que demonstram, por exemplo, que sistemas de redes neurais seriam capazes de identificar a orientação sexual das pessoas a partir de fotos de seu rosto (disponibilizadas em redes sociais), com alto grau de precisão, o que pode ter sérias repercussões jurídicas dependendo do contexto. Para maiores detalhes consultar: WANG, Yilun.; KOSINSKI, Michal. *Deep neural networks are more accurate than humans at detecting sexual orientation from facial images*. Journal of Personality and Social Psychology. American Psychological Association, 2017. Disponível em: <https://osf.io/zn79k/>. Acesso em 6 de junho de 2020.

<sup>27</sup> As linguagens naturais são aquelas desenvolvidas pelos seres humanos em seus sistemas de comunicação, e que se diferenciam das linguagens (artificiais) de programação de computadores, por exemplo. Atualmente se entende que, grosso modo, o conceito de língua natural compreende uma abstração, isto é, algo que se modifica dentro do tempo e do espaço, e que recebe consequentemente influências culturais e políticas. MATTHEWS, P. H. *Linguistics: a very short introduction*. Oxford: Oxford University Press, 2003. Capítulo 6.

complexidade do caso: processamento da sintaxe, da semântica, do contexto do discurso (a parte mais difícil de ser elaborada pelo computador), do raciocínio da conversa (com seu sequencial de mensagens inter-relacionadas), do planejamento da resposta e, finalmente, da produção da resposta de fato.<sup>28</sup>

Apesar das enormes dificuldades que envolvem o processamento de linguagem natural por um computador, os algoritmos utilizados nesta atividade têm se tornado cada vez mais sofisticados<sup>29</sup>, possibilitando que os diversos sistemas de inteligência artificial (como os indicados acima) adquiram uma aparência mais familiar aos seres humanos, facilitando seu uso pelo cidadão comum e disfarçando a complexidade de seu funcionamento.

Em síntese, nos parágrafos anteriores apresentou-se, de modo sucinto, o funcionamento de alguns tipos de sistemas de inteligência artificial que são comumente utilizados pelos cientistas da computação: o aprendizado de máquina, as redes neurais e o *deep learning*, além do processamento de linguagem natural. Assim, o conceito de inteligência artificial hoje engloba inúmeras técnicas que, embora não sejam exatamente recentes, podem ser facilmente obtidas no mercado até mesmo na forma de prestação de serviço especializado, ou de produto (imaterial) que pode ser adquirido prontamente para aplicação imediata tanto no setor público quando na iniciativa privada. Não obstante todas as divergências sobre a definição precisa do conceito de inteligência artificial, o seu mecanismo de operação é fundamentalmente o mesmo: algoritmos, sistemas de execução em cadeia de atos pré-programados que podem apresentar diversos níveis de complexidade, conforme se demonstra a seguir.

#### 2.4 Algoritmo e sistemas algorítmicos

O objetivo desta seção é trazer uma definição razoavelmente satisfatória acerca dos algoritmos (ou “programas de computador”) e seu funcionamento. Contudo, considerando o

---

<sup>28</sup> BATES, Madeleine. *Models of natural language understanding*. Proceedings of the National Academy of Sciences, v. 92, n. 22, 1995. p. 9977 e seguintes. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC40721/pdf/pnas01500-0075.pdf> Acesso em 23 de agosto de 2020.

<sup>29</sup> A ferramenta GPT-3 (*Generative Pre-trained Transformer 3*) da organização norte-americana OpenAI é um exemplo, que tem a capacidade de compreender instruções escritas e entregar os resultados desejados, aprendendo com o processo. O software reconhece imagens (oferecendo descrições) e diferentes tipos de estilos de textos (separando os e-mails relevantes de *spam*, por exemplo), sendo capaz de indicar o código de programação utilizado para realizar a tarefa. Em um teste, o GPT-3 obteve sucesso em gerar automaticamente uma carta de apresentação a partir de uma descrição de cargo e um currículo. Nesse sentido, este algoritmo atua com interpretação e geração de texto, funcionalidade que eventualmente poderá ser combinada com um mecanismo de geração de voz para tornar ainda mais difícil para um ser humano distinguir se conversa ao telefone com outra pessoa ou com um robô, o que traz inúmeros benefícios e riscos. A ferramenta pode ser consultada no seguinte endereço: <https://openai.com/blog/openai-api/>. Acesso em 23 de agosto de 2020.

propósito maior deste trabalho, não será o caso de exaurir o conceito ou de se aprofundar tecnicamente sobre este assunto, a exemplo do que foi feito no tópico anterior.

No contexto da ciência da computação, os programadores (humanos) são os profissionais responsáveis pelo desenho de fluxos lógicos<sup>30</sup> nos quais, a partir de dados de entrada determinados (informações inseridas pelo usuário do sistema, como por exemplo respostas de questionários dinâmicos ou ainda interpretação de texto feita por algoritmos), seja possível obter uma série de dados de saída desejados (o resultado da operação, informação condicionada e processada de acordo com os dados inicialmente inseridos na operação pelo usuário do sistema). Assim, os programadores fazem uso de paradigmas relacionados à construção do raciocínio que deverá ser utilizado pelo algoritmo, que, por sua vez, será estruturado de modo a resolver um problema objetivo.<sup>31</sup>

Assim, o algoritmo pode ser definido, grosso modo, como:<sup>32</sup>

[...] uma linguagem na qual se apresenta uma sequência de regras, declarações e operações que, aplicada sobre um conjunto de dados, nos permite solucionar problemas. A base de um algoritmo são as instruções à máquina de modo a executar tarefas específicas. [...] Um programa, ou um algoritmo, é uma sequência de instruções que pode ser executada por uma máquina. [...] Sendo assim, é importante que no processo de construção do algoritmo se elimine qualquer comando desnecessário ou incompreensível pela máquina. Conclusão, o computador é muito mais burro do que imaginamos. O computador possui um recurso bem limitado de instruções. [...] Um algoritmo, ou programa, é em essência isto: uma sequência de condicionais, repetições e comandos executáveis por uma máquina.

Nesse sentido, o algoritmo nada mais é do que um conjunto de comandos que são sistematicamente aplicados a determinados dados de entrada (*inputs*), para realizar uma tarefa específica em um número finito de passos delimitados. O algoritmo, portanto, pode ser compreendido como uma fórmula ou ferramenta para execução de uma operação, sem ter “consciência” sobre seu funcionamento. Aqui, os programadores humanos exercem o papel de redigir, na linguagem de programação selecionada, os comandos para o computador, que executará a atividade conforme descrito. Devido às limitações da máquina, que não possui

---

<sup>30</sup> A Teoria dos Grafos possui grande importância para a estruturação desses fluxos lógicos, sendo especialmente valiosa na resolução de problemas das áreas de Ciência e Engenharia da Computação. Sobre este assunto, ver: PRESTES, Edson. *Teoria dos grafos*. Porto Alegre: 2011, vol. 4, 2016. Disponível em: <http://www.inf.ufrgs.br/~prestes/Courses/Graph%20Theory/Livro/ParteLivroGrafos.pdf>. Acesso em 16 de março de 2020.

<sup>31</sup> MACHADO, Anderson Fraiha; MENDES, Leilani Dian; GARCEZ, Lucas Nogueira. *Introdução à Arquitetura e Engenharia Jurídica com Lawtex*. São Paulo: Looplex, 2018.

<sup>32</sup> MACHADO, Anderson Fraiha; MENDES, Leilani Dian; GARCEZ, Lucas Nogueira. *Idem*, *ibidem*. pp. 90 - 95

contexto (como os humanos), a operação deve ser minuciosamente descrita e segmentada em cada passo, como se fossem pequenos problemas sucessivos, até que se atinja o resultado final. A figura 1, a seguir, representa uma versão simples de um diagrama (também chamado de “árvore de decisão”) onde estão representados todos os passos e decisões que um indivíduo hipotético poderia considerar em seu processo de decisão sobre passear no parque de seu bairro. No desenho, os losangos representam etapas condicionais que dependem dos dados inseridos no sistema (como uma resposta de “sim” ou “não” pelo usuário) para que o algoritmo chegue ao resultado, representado por elipses. As diversas linguagens de programação possuem a função de traduzir esse esquema para o computador, possibilitando a execução das respectivas instruções pelas máquinas.

Os comandos representados na linguagem de programação designada se materializam em milhares de pequenas correntes elétricas que percorrem o processador (o “cérebro” do computador), para entregar o resultado final de acordo com o que foi descrito pelo programador. Assim, é possível fazer com que o computador funcione, permitindo ao usuário utilizar redes sociais, assistir filmes, jogar *videogames*, ler e editar um documento digitalizado, etc. Tudo isso consiste essencialmente no resultado do funcionamento dos algoritmos.

Atualmente os algoritmos se converteram em uma alternativa bastante valiosa para automatizar um procedimento, por meio de elementos fundamentais da própria composição algorítmica, que expressa o resultado de sua operação por meio de um protocolo formal com alto nível de abstração. Sua complexidade pode variar bastante, mas sua estrutura se mantém fundamentada em cadeias de ação pré-programadas cujo objetivo é a resolução de um problema tangível, com base nos dados disponibilizados inicialmente.

Conforme mencionado anteriormente a respeito do aprendizado de máquina, os sistemas algorítmicos complexos fazem uso de uma análise estatística dos dados disponíveis, de modo que quanto mais sofisticado for o algoritmo, maior será a sua capacidade de processamento das informações, o que deixará o sistema mais rápido e preciso. Entretanto, como ocorre em qualquer análise estatística, o bom resultado de uma operação algorítmica estará sempre condicionado à sua construção lógica e sobretudo à qualidade e à quantidade das informações disponibilizadas em seu estágio inicial, a partir das quais o algoritmo irá criar um modelo para identificar padrões aplicáveis a um novo conjunto de informações.<sup>33</sup>

---

<sup>33</sup> SURDEN, Harry. *Machine Learning and Law*. Idem, ibidem. p. 93.

Figura 1: Exemplo de árvore de decisão.



Fonte: Looplex, 2018.

Deste modo, apesar de todo o entusiasmo atual a respeito das potencialidades dos sistemas de inteligência artificial alimentados por algoritmos sofisticados, é inegável que o

programador humano exerce um papel fundamental<sup>34</sup> sobre as informações que serão fornecidas ao sistema algorítmico, além da maneira pela qual o computador irá tratá-las. Este detalhe se torna essencial em relação ao resultado da operação automatizada, isto é, a solução que o algoritmo apresenta ao final do processamento dos dados pode variar substancialmente na medida em que haja uma seleção diversa das informações iniciais.

Assim, este capítulo teve como objetivo específico realizar breve exposição a respeito dos conceitos principais envolvendo a ciência da computação, sem grande aprofundamento técnico, incluindo alguns exemplos de operação dos sistemas mais difundidos de inteligência artificial. Também se abordou o funcionamento específico dos algoritmos que constituem o mecanismo central de operação dos computadores. No próximo capítulo será realizada breve contextualização da pesquisa, inclusive apresentando determinadas considerações a respeito da legislação que abrange o tema atualmente.

---

<sup>34</sup> O reconhecimento da relevância que o papel do programador de computadores exerce no funcionamento do algoritmo levou algumas associações internacionais de programadores e engenheiros de computação, como a *Association for Computing Machinery* (ACM) e o *Institute for Electrical and Electronic Engineers* (IEEE) a desenvolver um código de ética para evitar alguns dos problemas que a tecnologia pode trazer, inclusive relacionados ao viés algorítmico.

### 3 Contexto da pesquisa

Este capítulo tem por objetivo específico realizar breve contextualização a respeito do tema desta pesquisa, apresentando considerações sucintas sobre as implicações jurídicas das tecnologias de inteligência artificial e as respectivas normas aplicáveis, incluindo a Constituição. Em seguida, o texto expõe resumidamente alguns pontos principais do debate legislativo sobre o tema.

Em relação ao ordenamento jurídico brasileiro, a Constituição de 1988 resguarda uma série de garantias para que os indivíduos exerçam seus direitos e deveres de modo seguro e previsível, embora seja fundamental que quase sempre a legislação infraconstitucional necessite intervir para detalhar os mandamentos constitucionais e regular, na prática, as disposições da Lei Maior. É por essa via que as diversas instituições estatais podem implementar as políticas públicas previstas em lei para concretizar as metas estabelecidas pelo texto constitucional.

De início, o Marco Civil da Internet (Lei Nº 12.965, de 23 de abril de 2014)<sup>35</sup> representou a mais importante norma voltada especificamente para a regulação da Internet no Brasil, capaz de servir de orientação para futuras políticas públicas a respeito deste tema. Com sua redação de forte caráter principiológico (como a neutralidade da rede, a liberdade de expressão e privacidade dos usuários), o Marco Civil se tornou referência mundial para a regulação da Internet, uma área que permanece em constante transformação. Em seu texto se destacam as disposições que tratam da responsabilidade civil de provedores em relação ao conteúdo criado e/ou publicado por terceiros (incluindo também dispositivos que apresentam relevância para a esfera penal, relacionados a investigações de crimes cibernéticos), além de estimular debates sobre o direito ao esquecimento dentro do ordenamento jurídico brasileiro.

O Marco Civil da Internet também serviu inicialmente como um mecanismo geral para a regulação dos dados pessoais, uma vez que até 2018 o Brasil não dispunha de lei especial sobre o assunto. Deste modo, as disposições normativas instituídas pelo Marco Civil “inauguravam o tratamento da tutela de dados no que diz respeito à rede no País”, pois “até então se contava apenas com dispositivos ora muito genéricos, como o constante do artigo 21

---

<sup>35</sup> Disponível em: [http://www.planalto.gov.br/ccivil\\_03/\\_ato2011-2014/2014/lei/112965.htm](http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/112965.htm). Acesso em 17 de agosto de 2020.

do Código Civil<sup>36</sup>, ora muito setoriais, como o artigo 43 do Código de Defesa do Consumidor<sup>37</sup>, como afirmam os professores Carlos Affonso Souza e Ronaldo Lemos.<sup>38</sup>

É exatamente neste ponto que a entrada em vigor da Lei Geral de Proteção de Dados<sup>39</sup>, comumente chamada de “LGPD”, representou a mais importante iniciativa brasileira de regulamentação do uso de dados pessoais que consiste na “matéria-prima” fundamental do chamado *big data*, a nova ciência dos dados em massa que possibilita o desenvolvimento de mecanismos cada vez mais complexos de inteligência artificial, como os abordados no capítulo anterior. Aqui o legislador brasileiro em princípio buscou como referência principal o Regulamento Geral Europeu sobre a Proteção de Dados<sup>40</sup> (RGPD), promulgado pelo Parlamento da União Europeia em 2016.

O cenário tecnológico atual aponta para o aumento progressivo no volume (astronômico) de dados que é produzido a partir dos diversos aparelhos eletrônicos interconectados que se tornaram parte da vida cotidiana em quase todos os países nas últimas décadas<sup>41</sup>. Nestas condições, a LGPD terá impacto no cotidiano dos cidadãos brasileiros em inúmeras situações<sup>42</sup>, desde as mais relevantes até aquelas triviais, como solicitações de crédito

---

<sup>36</sup> Código Civil, Art. 21. A vida privada da pessoa natural é inviolável, e o juiz, a requerimento do interessado, adotará as providências necessárias para impedir ou fazer cessar ato contrário a esta norma.

<sup>37</sup> Código de Defesa do Consumidor, Art. 43. O consumidor, sem prejuízo do disposto no art. 86, terá acesso às informações existentes em cadastros, fichas, registros e dados pessoais e de consumo arquivados sobre ele, bem como sobre as suas respectivas fontes. § 1º Os cadastros e dados de consumidores devem ser objetivos, claros, verdadeiros e em linguagem de fácil compreensão, não podendo conter informações negativas referentes a período superior a cinco anos. § 2º A abertura de cadastro, ficha, registro e dados pessoais e de consumo deverá ser comunicada por escrito ao consumidor, quando não solicitada por ele. § 3º O consumidor, sempre que encontrar inexatidão nos seus dados e cadastros, poderá exigir sua imediata correção, devendo o arquivista, no prazo de cinco dias úteis, comunicar a alteração aos eventuais destinatários das informações incorretas. § 4º Os bancos de dados e cadastros relativos a consumidores, os serviços de proteção ao crédito e congêneres são considerados entidades de caráter público. § 5º Consumada a prescrição relativa à cobrança de débitos do consumidor, não serão fornecidas, pelos respectivos Sistemas de Proteção ao Crédito, quaisquer informações que possam impedir ou dificultar novo acesso ao crédito junto aos fornecedores. § 6º Todas as informações de que trata o caput deste artigo devem ser disponibilizadas em formatos acessíveis, inclusive para a pessoa com deficiência, mediante solicitação do consumidor.

<sup>38</sup> SOUZA, Carlos Affonso Pereira de; SILVA JUNIOR, Ronaldo Lemos da. *Marco Civil da Internet: construção e aplicação*. 1. ed. Juiz de Fora: Editar, 2016. v. 1. p. 25. Disponível em: [https://itsrio.org/wp-content/uploads/2017/02/marco\\_civil\\_construcao\\_aplicacao.pdf](https://itsrio.org/wp-content/uploads/2017/02/marco_civil_construcao_aplicacao.pdf). Acesso em 18 de agosto de 2020.

<sup>39</sup> Lei nº 13.709, sancionada em 14 de agosto de 2018 pelo então presidente Michel Temer e vigente desde 18 de setembro de 2020. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/\\_ato2015-2018/2018/lei/L13709.htm](http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm) (acesso em 21/02/2020).

<sup>40</sup> Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:32016R0679&from=PT#d1e40-1-1>. Acesso em 16 de março de 2020.

<sup>41</sup> Disponível em: <http://www.ihu.unisinos.br/78-noticias/593456-o-big-data-apresenta-uma-multimetodologia-entrevista-com-walter-sosa-escudero>. Acesso em 22 de fevereiro de 2020.

<sup>42</sup> O Direito brasileiro buscou se adequar a essa nova conjuntura social, com o objetivo de estabelecer regras mais específicas para a coleta, armazenamento e uso de informações pessoais por instituições públicas e privadas. Por exemplo, existe a prática comum de solicitar do número do CPF do consumidor na aquisição de produtos e serviços. Para maiores detalhes, consultar: <https://www.nexojournal.com.br/expresso/2020/02/15/O-que-est%C3%A1-em-jogo-quando-voc%C3%AA-d%C3%A1-seu-CPF-na-hora-da->

junto a instituições financeiras ou simples pedidos de refeições, por exemplo, na medida em que se consolida o papel do *big data* na intermediação de praticamente todos os aspectos da vida social contemporânea.

Entretanto, diferentemente da norma europeia, a lei brasileira<sup>43</sup> deixou de prever expressamente a obrigatoriedade de revisão humana sobre uma operação algorítmica. Ou seja, não há obrigação de que o resultado do tratamento automatizado de dados seja analisado por um operador humano, mesmo quando o usuário do serviço se sinta prejudicado. Embora seja bastante improvável que os computadores excedam e mesmo substituam por completo a força de trabalho humana nas próximas décadas, parece inevitável a tendência de que as máquinas continuem ganhando relevância em um número cada vez maior de atividades.

Deste modo, a discussão acadêmica e política em âmbito internacional há alguns anos vem alertando a sociedade para os possíveis efeitos negativos da atividade algorítmica sobre a esfera individual de interesses jurídicos. Neste sentido, este trabalho pretende analisar se a sociedade brasileira dispõe de meios jurídicos suficientes que ofereçam uma resposta institucional à altura dos desafios desta nova realidade, na qual o avanço da ciência de dados representa, simultaneamente, possibilidade de melhorias e fonte de problemas aos usuários.<sup>44</sup>

Conforme indicado na introdução, o foco desta pesquisa está no chamado “viés algorítmico”, que trata das decisões tomadas por sistemas de algoritmos que podem estar comprometidas por diversos tipos de vieses, acarretando prejuízos aos indivíduos que se encontram sujeitos aos resultados dessas operações eletrônicas automatizadas. Esse tema se insere em uma conjuntura mais ampla, que trata dos impactos legais causados pelas aplicações da ciência de dados em massa, sobretudo no que diz respeito aos “direitos democráticos e a privacidade”.<sup>45</sup>

Este cenário envolve não apenas a relação entre a LGPD (ou outra lei específica) e a discriminação algorítmica<sup>46</sup>, mas, considerando seus efeitos na sociedade como um todo, é

---

[compra?utm\\_medium=Social&utm\\_campaign=Echobox&utm\\_source=Twitter#Echobox=1582314608](https://www.conjur.com.br/2019-jul-24/renan-lopergolo-lacunas-lgpd-casos-faceapp). Acesso em 22 de fevereiro de 2020.

<sup>43</sup> Como é possível constatar a partir da leitura dos artigos 20 e 22 das leis brasileira e europeia, respectivamente.

<sup>44</sup> Disponível em: <https://www.conjur.com.br/2019-jul-24/renan-lopergolo-lacunas-lgpd-casos-faceapp>. Acesso em 12 de fevereiro de 2020.

<sup>45</sup> MOLNAR, Adam. *Technology, law, and the formation of (il)liberal democracy?*, Surveillance and society, vol. 15, no. 3/4, 2017, pp. 318-388. Disponível em <https://ojs.library.queensu.ca/index.php/surveillance-and-society/article/view/6645>. Acesso em 3 de junho de 2020. Aqui o autor chama a atenção para o papel relevante exercido pelos direitos fundamentais elencados na constituição de cada Estado no sentido de resguardar os indivíduos de investidas autoritárias por parte das autoridades públicas devido ao mau uso do *big data*, o que também indica a responsabilidade de juízes e legisladores no sentido de efetivar tais garantias na prática.

<sup>46</sup> O termo “discriminação” aqui é utilizado em um sentido amplo para se referir a impactos prejudiciais descomunais - intencionais ou não – em que o desenho, a implementação e a utilização de sistemas algorítmicos podem causar em determinados grupos sociais ou indivíduos que compartilham certas características. Esta

preciso analisar a capacidade do ordenamento jurídico brasileiro em solucionar os desafios que já estão sendo propostos por essa nova condição. Este caso pode até mesmo gerar um debate envolvendo conflitos de direitos fundamentais. Por exemplo, a liberdade econômica (das empresas em geral, especialmente as de tecnologia) versus a não-discriminação (igualdade), como alguns dos recentes argumentos utilizados nas discussões legislativas permitem concluir, conforme se demonstra adiante.

Neste sentido, cabe avaliar se os dispositivos existentes na legislação (como exemplifica o Artigo 20<sup>47</sup> da LGPD), são suficientes para oferecer uma solução satisfatória ao problema da discriminação algorítmica<sup>48</sup> e garantir o cumprimento das normas constitucionais que protegem os direitos fundamentais dos cidadãos, o que deve se tornar um desafio ainda mais presente na medida em que se intensifica a mudança de paradigma tecnológico atualmente em curso. Conforme indicado anteriormente, este assunto vem recebendo cada vez mais atenção nas pesquisas acadêmicas feitas na área jurídica em âmbito internacional, e se coloca como um grande debate que precisa ser realizado pela sociedade brasileira nos próximos anos, de modo a evitar eventuais prejuízos à garantia dos direitos humanos de nossa população.

Como se demonstra adiante, este tema é tão complexo quanto urgente para o Direito, pois a própria geração massiva de dados torna impossível seu manejo sem a utilização das próprias ferramentas proporcionadas pelo *big data*. Este assunto merece ainda mais atenção por parte de todos os agentes sociais na medida em que tecnologias nascentes (como a conexão de alta velocidade 5G e a chamada Internet das Coisas<sup>49</sup>) se tornarem a nova realidade mundial

---

pesquisa não tem por objetivo propor uma conclusão legal sobre a discriminação a partir da legislação penal ou de qualquer outra. Alguns casos de discriminação podem ser não intencionais e até imprevistos, outros podem ser o resultado de uma decisão política deliberada de concentrar serviços ou assistência naqueles que mais necessitam, como no caso das ações afirmativas. Outros ainda podem criar consequências adversas para populações específicas que criam ou exacerbam a desigualdade de oportunidades para segmentos sociais ou cidadãos que já estão em desvantagem. É neste último caso que o trabalho se concentra, especificamente.

<sup>47</sup>Lei nº 13.709/2018, Art. 20: O titular dos dados tem direito a solicitar a revisão de decisões tomadas unicamente com base em tratamento automatizado de dados pessoais que afetem seus interesses, incluídas as decisões destinadas a definir o seu perfil pessoal, profissional, de consumo e de crédito ou os aspectos de sua personalidade. § 1º O controlador deverá fornecer, sempre que solicitadas, informações claras e adequadas a respeito dos critérios e dos procedimentos utilizados para a decisão automatizada, observados os segredos comercial e industrial. § 2º Em caso de não oferecimento de informações de que trata o § 1º deste artigo baseado na observância de segredo comercial e industrial, a autoridade nacional poderá realizar auditoria para verificação de aspectos discriminatórios em tratamento automatizado de dados pessoais.

<sup>48</sup>Especificamente sobre a LGPD, existem diversas questões que mereceriam uma investigação própria, como os efeitos do Art. 20 da LGPD no tratamento automatizado de dados pessoais ou ainda se seria possível que falta de revisão humana por si já representasse um prejuízo em potencial para os usuários.

<sup>49</sup>Maiores detalhes sobre a tecnologia 5G e a Internet das Coisas disponíveis em

<https://agenciabrasil.ebc.com.br/geral/noticia/2020-03/agencia-brasil-explica-o-que-e-tecnologia-5g>. Acesso em 3 de junho de 2020.

nos próximos anos<sup>50</sup>, trazendo consigo outras mudanças significativas para nosso modo de vida atual, e com isso diversas questões para quais a sociedade certamente esperará respostas do Direito.

Considerando o caso brasileiro, o texto constitucional promulgado em 1988 dispôs em seu artigo 3º<sup>51</sup> que um dos objetivos fundamentais da República é combater a discriminação em suas diversas formas. Trata-se de uma meta com força vinculante que deve ser aplicada em todas as decisões e atividades desenvolvidas por cada integrante do Estado brasileiro, inclusive durante o exercício de fiscalização da sociedade civil. Entretanto, assim como ocorre com as demais metas constitucionais, é quase sempre necessário que uma política pública específica seja formulada pelo legislador infraconstitucional e executada pelo poder executivo para que as previsões da Constituição sejam efetivadas na prática.

Neste sentido, a atuação positiva do poder público para combater a discriminação é imprescindível, de acordo com Philip Gil França, “ao desenvolvimento intersubjetivo dos partícipes do sistema constitucional”, isto é, “quando o mínimo possível de viabilidade deste desiderato é sentido na vida daqueles que estão sob a égide de sua regulação”.<sup>52</sup> Na visão do professor, a política pública pode ser definida como “a organização de esforços estatais para alcançar um claro objetivo predeterminado, mediante um planejado caminho, a partir da demonstração objetiva de realização de um interesse público constitucionalmente previsto”.<sup>53</sup>

Deste modo, conforme mencionado acima, o Marco Civil da Internet constituiu a primeira tentativa do legislador de balizar as futuras políticas públicas para regular o mundo digital. Embora não trate apenas do uso de dados pessoais nem estabeleça regras específicas de funcionamento para algoritmos e sistemas de informação, sua redação teve o mérito especial de aproximar a sociedade das discussões sobre a regulação da Internet, além de sedimentar o alcance da responsabilidade civil para os provedores de Internet (positivando o entendimento jurisprudencial de tribunais nacionais e estrangeiros no sentido de não responsabilizar o

---

<sup>50</sup> A implementação da tecnologia 5G ao redor do mundo já tem contribuído para acirrar a disputa geopolítica entre Estados Unidos e China. Washington vem seguidamente acusando o país asiático de obter vantagens indevidas incluindo a possibilidade de controle sobre a comunicação internacional caso a infraestrutura implantada seja desenvolvida pela empresa chinesa Huawei, o que é negado por Pequim. A notícia a seguir (em inglês) exemplifica a proporção que esta controvérsia pode tomar: <https://www.theguardian.com/technology/2020/jun/02/us-senator-huawei-5g-is-like-soviets-building-wests-submarines>. Acesso em 3 de junho de 2020.

<sup>51</sup> CRFB/88, Art. 3º: Constituem objetivos fundamentais da República Federativa do Brasil: (...) IV - promover o bem de todos, sem preconceitos de origem, raça, sexo, cor, idade e quaisquer outras formas de discriminação.

<sup>52</sup> FRANÇA, Phillip Gil. *Objetivos Fundamentais da República, Escolhas Públicas e Políticas Públicas: Caminhos de Concretização dos Benefícios Sociais Constitucionais*. Direitos sociais e políticas públicas I. 1ed. Curitiba: Clássica Editora, 2014, v. 25, p. 230-244.

<sup>53</sup> FRANÇA, Phillip Gil. *Idem*, ibidem. p. 236.

provedor da conexão pelas condutas de seus usuários).<sup>54</sup> Por outro lado, como não poderia deixar de ser em uma democracia, existem críticas<sup>55</sup> ao fato desta lei privilegiar fortemente a liberdade de expressão a ponto de ser preciso recorrer ao Judiciário para retirar de circulação eventual conteúdo que viole por exemplo a intimidade, privacidade ou a proteção ao consumidor. De todo modo, o Marco Civil não será analisado nesta pesquisa por não reger especificamente os casos envolvendo violações de direitos causados por vieses algorítmicos, ainda que seja válida a discussão sobre o papel que o Marco Civil possa desempenhar enquanto marco regulatório geral para o funcionamento da Internet brasileira, o que pode ser um tema de investigação interessante em pesquisas futuras.

Já o Art. 20 da LGPD possibilita ao interessado solicitar a revisão de decisões tomadas exclusivamente por algoritmos (sem intervenção humana). Sobre este ponto houve uma iniciativa política do senador Styvenson Valentim (vinculado ao partido político Podemos, do Rio Grande do Norte), que apresentou o Projeto de Lei 4.496/2019 com o objetivo de alterar a LGPD no que diz respeito à definição da expressão “decisão automatizada”. Segundo o autor, este projeto partiu da perspectiva de existirem diversas maneiras de realizar as “decisões automatizadas” previstas na LGPD, sendo necessário indicar expressamente os mecanismos de aprendizado de máquina e inteligência artificial no texto da lei. Essa ação serviria para garantir transparência aos cidadãos, na medida em que facilitaria o acesso a “informações claras e adequadas a respeito dos critérios e dos procedimentos utilizados para a decisão automatizada”, conforme previsão do §1º do Art. 20 da LGPD. Segundo o senador, “os responsáveis pelo tratamento de dados não fornecem esclarecimentos apropriados para decisões baseadas em técnicas de inteligência artificial ou de outras igualmente complexas”.<sup>56</sup>

Como indicado anteriormente, a lei brasileira se inspirou no modelo adotado pela União Europeia para regular este tema. Entretanto, o texto da norma editada pelo parlamento europeu contém explicitamente a possibilidade de revisão humana como uma alternativa aos interessados na reavaliação das referidas decisões<sup>57</sup>, o que foi vetado pela Presidência da

---

<sup>54</sup> SOUZA, Carlos Affonso Pereira de; SILVA JUNIOR, Ronaldo Lemos da. *Idem*, *ibidem*. p. 98.

<sup>55</sup> CARVALHO, P. H.. *O Marco Civil da Internet: Uma análise sobre a constitucionalidade do artigo 19*. Revista da Faculdade de Direito do Sul de Minas. Pouso Alegre, 2017, v. 33, p. 228-244. Disponível em: <https://www.fdsu.edu.br/adm/artigos/6917c36392274e9b6393c7f7a7bddd1.pdf>. Acesso em 18 de agosto de 2020.

<sup>56</sup> Maiores detalhes disponíveis em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/138136>. Acesso em 12 de agosto de 2020.

<sup>57</sup> Regulamento (UE) 2016/679 do Parlamento Europeu e do Conselho, Artigo 22: Decisões individuais automatizadas, incluindo definição de perfis 1. O titular dos dados tem o direito de não ficar sujeito a nenhuma decisão tomada exclusivamente com base no tratamento automatizado, incluindo a definição de perfis, que produza efeitos na sua esfera jurídica ou que o afete significativamente de forma similar. 2. O nº 1 não se aplica se a decisão: a) for necessária para a celebração ou a execução de um contrato entre o titular dos dados e um

República na versão brasileira. Essa possibilidade foi inicialmente adotada pelo legislador brasileiro, contudo após o veto presidencial referente ao parágrafo terceiro do diploma legal, a eventual nova análise que seja solicitada pelo interessado não será obrigatoriamente realizada por um ser humano.

As razões do veto presidencial alegaram a inviabilização dos modelos de negócios de empresas privadas, o que pode parecer algo questionável quando se observa que a economia europeia é bastante dependente do setor de serviços (característica que compartilha com a economia brasileira) e ainda assim determinou a intervenção humana neste cenário. Em especial, a norma brasileira pode se tornar problemática sobretudo por condicionar a preservação de direitos fundamentais<sup>58</sup> a interesses econômicos privados. Entretanto, até o momento não há evidências de que a regulação europeia resultou na inviabilização de negócios ou prejuízos econômicos para as empresas que atuam nos países que integram a União Europeia. Além disso, há o risco de consistir em medida provavelmente inútil quanto ao que se propõe e talvez mais custosa para a sociedade face à possibilidade de sobrecarga do poder judiciário<sup>59</sup> com demandas que poderiam ser sanadas administrativamente entre as partes. A mensagem da Presidência da República nº 288/2019, dirigida à Presidência do Senado Federal, afirma que

a propositura legislativa, ao dispor que toda e qualquer decisão baseada unicamente no tratamento automatizado seja suscetível de revisão humana, contraria o interesse público, tendo em vista que tal exigência inviabilizará os modelos atuais de planos de negócios de muitas empresas, notadamente das *startups*, bem como impacta na análise de risco de crédito e de novos modelos de negócios de instituições financeiras, gerando efeito negativo na oferta de crédito aos consumidores, tanto no que diz respeito à qualidade das garantias, ao volume de crédito contratado e à composição de

---

responsável pelo tratamento; b) for autorizada pelo direito da União ou do Estado-Membro a que o responsável pelo tratamento estiver sujeito, e na qual estejam igualmente previstas medidas adequadas para salvaguardar os direitos e liberdades e os legítimos interesses do titular dos dados; ou c) For baseada no consentimento explícito do titular dos dados. 3. Nos casos a que se referem o nº 2, alíneas a) e c), o responsável pelo tratamento aplica medidas adequadas para salvaguardar os direitos e liberdades e legítimos interesses do titular dos dados, designadamente o direito de, pelo menos, obter intervenção humana por parte do responsável, manifestar o seu ponto de vista e contestar a decisão. 4. As decisões a que se refere o nº 2 não se baseiam nas categorias especiais de dados pessoais a que se refere o artigo 9º, nº 1, a não ser que o nº 2, alínea a) ou g), do mesmo artigo sejam aplicáveis e sejam aplicadas medidas adequadas para salvaguardar os direitos e liberdades e os legítimos interesses do titular.

<sup>58</sup> A liberdade empresarial é um direito fundamental (resguardado no inciso IV do Art. 1º da Constituição de 1988), de modo que o veto presidencial à obrigatoriedade de revisão humana em decisões automatizadas poderia talvez ser compreendido como um exemplo que enseja uma colisão entre direitos fundamentais. Esta colisão “deve ser enfrentada somente no momento da possível justificação de uma intervenção estatal(...) porque um direito fundamental de outro titular de direito pode estar limitando o exercício do direito fundamental atingido pela medida ou omissão estatal.” DIMOULIS, Dimitri.; MARTINS, Leonardo. *Teoria geral dos direitos fundamentais*. 5. ed. São Paulo: Atlas, 2014. p. 169.

<sup>59</sup> Em consulta à jurisprudência dos Tribunais de Justiça dos estados do Rio de Janeiro, Rio Grande do Sul e de São Paulo, além do Superior Tribunal de Justiça e do Supremo Tribunal Federal não foram identificados julgados relativos à discriminação algorítmica, especificamente.

preços, com reflexos, ainda, nos índices de inflação e na condução da política monetária.

Assim, a redação final da lei dificultou que decisões automatizadas (isto é, tomadas por algoritmos) sejam corrigidas por meio de intervenção humana e deixou de ser ferramenta valiosa no combate a práticas discriminatórias (que são abordadas no capítulo 4), como, por exemplo, eventual decisão que recuse a oferta de crédito a cliente de instituição financeira (ou que considere que o risco do mútuo seja elevado tornando a taxa de juros mais onerosa, conseqüentemente) em decorrência de características raciais e/ou étnicas do solicitante ou de seu endereço residencial, mesmo sob prova documental que comprove aptidão financeira. Neste caso, não se discute que a concessão de crédito é inerente à liberdade negocial das instituições financeiras, entretanto a falta de transparência no tratamento automatizado dos dados impede a verificação de possíveis critérios discriminatórios na operação do algoritmo. Novamente, o legislador não inseriu padrões no texto da LGPD para balizar objetivamente o tratamento de dados e garantir segurança jurídica a todas as partes. Entretanto, no âmbito judicial o artigo 7º da Resolução 332/2020 do Conselho Nacional de Justiça tomou por base a LGPD para estabelecer medidas de combate à discriminação no uso de ferramentas de inteligência artificial pelo Poder Judiciário.<sup>60</sup> Membros da academia<sup>61</sup> também tem defendido que a proteção aos direitos individuais consagrados pela LGPD não representa óbice à livre iniciativa empresarial; pelo contrário, possibilitaria maior segurança jurídica e engajamento social capaz de fomentar o ambiente econômico.

Diante deste cenário em que se identificam inúmeros desafios representados pelo avanço na utilização de sistemas de inteligência artificial e suas conseqüências sociais e jurídicas, é possível levantar a hipótese de que a discriminação promovida por algoritmos não encontra

---

<sup>60</sup> Res. 332/2020 CNJ, Art. 7º: As decisões judiciais apoiadas em ferramentas de Inteligência Artificial devem preservar a igualdade, a não discriminação, a pluralidade e a solidariedade, auxiliando no julgamento justo, com criação de condições que visem eliminar ou minimizar a opressão, a marginalização do ser humano e os erros de julgamento decorrentes de preconceitos. § 1º Antes de ser colocado em produção, o modelo de Inteligência Artificial deverá ser homologado de forma a identificar se preconceitos ou generalizações influenciaram seu desenvolvimento, acarretando tendências discriminatórias no seu funcionamento. § 2º Verificado viés discriminatório de qualquer natureza ou incompatibilidade do modelo de Inteligência Artificial com os princípios previstos nesta Resolução, deverão ser adotadas medidas corretivas. § 3º A impossibilidade de eliminação do viés discriminatório do modelo de Inteligência Artificial implicará na descontinuidade de sua utilização, com o consequente registro de seu projeto e as razões que levaram a tal decisão. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3429>. Acesso em 29 de novembro de 2020.

<sup>61</sup> Autores como como Giovana Carloni, vinculada ao Centre for Information Policy Leadership (CIPL), Carlos Affonso Souza, membro do Instituto de Tecnologia e Sociedade do Rio de Janeiro (ITS-Rio) e Bruno Ricardo Bioni, integrante da organização *think tank* Data Privacy Brasil, que participaram de um debate virtual a respeito das implicações das medidas de combate à pandemia de Covid-19 em face da LGPD. A íntegra está disponível no seguinte endereço eletrônico: <https://youtu.be/oBjQAfaf-xc>. Acesso em 8 de abril de 2020.

resposta específica por parte da legislação brasileira, o que representa uma lacuna regulatória sobretudo em perspectiva do Art. 3º da Constituição brasileira.

No âmbito da LGPD, por exemplo, a redação atual poderá ser insuficiente para coibir práticas discriminatórias por não prever a obrigatoriedade de intervenção humana na revisão de decisões automatizadas. Também cabe refletir sobre as condições da Autoridade Nacional de Proteção de Dados quanto à resposta, de modo solitário, de todas as demandas em potencial envolvendo casos de discriminação algorítmica. Embora existam outros dispositivos legais (sem falar na própria aplicação direta das normas constitucionais<sup>62</sup>) que protejam os indivíduos de situações discriminatórias, a LGPD poderia representar a principal iniciativa do poder público no que diz respeito à proteção dos direitos individuais e coletivos no cenário socioeconômico atual, onde as operações automatizadas efetuadas por algoritmos têm um impacto cada vez maior na esfera de interesses jurídicos dos cidadãos.

Assim, este capítulo teve por objetivo específico expor o contexto tecnológico e jurídico que abrange esta pesquisa, considerando as normas aplicáveis e o debate legislativo sobre o tema. No próximo capítulo, o foco consiste em apresentar e analisar alguns dos diversos tipos de vieses que podem ocorrer no funcionamento “normal” dos algoritmos, e que podem alimentar e ampliar diversas situações discriminatórias.

---

<sup>62</sup> Seguindo a classificação tripartite proposta por José Afonso da Silva, as normas constitucionais de eficácia plena podem constituir a principal garantia nos casos envolvendo a discriminação algorítmica, como será abordado adiante. A doutrina inclui o Art. 3º, CRFB/88 entre as normas de eficácia limitada (de princípio programático), que tem por objetivo vincular a atuação do legislador quando da elaboração de uma norma infraconstitucional. DUTRA, Luciano. *Direito Constitucional Essencial*. Rio de Janeiro: Forense. 2ªed. 2016. p.48.

#### 4 Viés algorítmico e discriminação

Conforme indicado anteriormente, toda inovação científica traz consigo o potencial de servir ao ser humano de modo a facilitar a execução de alguma tarefa ou resolver determinado problema. Novas ferramentas tecnológicas, neste sentido, podem trazer benefícios ou riscos para a sociedade, dependendo da forma como são utilizadas. Um exemplo clássico é a energia nuclear, que pode fornecer eletricidade de maneira menos nociva ao meio ambiente (sobretudo em comparação aos combustíveis fósseis), mas que também pode representar enorme potencial destrutivo quando utilizada para fins bélicos. Cabe à sociedade decidir o modo pelo qual a nova tecnologia será empregada.

Na medida em que as ferramentas de inteligência artificial são adotadas progressivamente pela sociedade, torna-se perceptível algumas consequências indesejadas. Sobretudo, a aplicação de sistemas de inteligência artificial para tomar decisões em substituição a agentes humanos e instituições leva a preocupações sobre como garantir justiça, imparcialidade e responsabilidade nos resultados dessas operações automatizadas. Concretamente, a inteligência artificial utilizada para controlar equipamentos do mundo real causa preocupações com a segurança, especialmente porque os sistemas estão expostos a toda a complexidade do ambiente humano.<sup>63</sup> Mitigar esses riscos requer o reconhecimento de que a solução deve incorporar os aspectos jurídicos e técnicos (da própria computação), considerando que justiça e segurança estão relacionadas. Como se demonstrará no próximo capítulo, a abordagem deve ser interdisciplinar, tanto por parte de empresas quanto do poder público, no esforço de evitar a discriminação algorítmica intencional, evitando resultados prejudiciais indesejados e gerando as evidências necessárias para dar às partes interessadas a confiança fundamentada de que as falhas não intencionais são raras e improváveis, ou seja, de que o uso da ferramenta de inteligência artificial proposta é seguro.

Neste momento, entretanto, o objetivo específico desta terceira seção do texto é apresentar e analisar os principais tipos de vieses que originam situações discriminatórias em mecanismos de inteligência artificial, além de realizar breve exposição do próprio conceito jurídico de discriminação. Assim, este capítulo realiza uma exposição sobre o viés algorítmico

---

<sup>63</sup> Um exemplo grave diz respeito aos carros autônomos que apresentam dificuldade de reconhecer as pessoas com tons de pele escuros como seres humanos, o que coloca esses indivíduos sob maior risco de acidentes. Embora carros autônomos sejam uma realidade ainda distante para a maioria das pessoas, essa falha poderia gerar acidentes graves nos casos dos mecanismos de freio emergencial que estão sendo adotados nos modelos mais recentes de automóveis. Para mais detalhes sobre viés racial em carros autônomos, consultar: WILSON, Benjamin; HOFFMAN, Judy; MORGENSTERN, Jamie. *Predictive Inequity in Object Detection*. arXiv preprint arXiv:1902.11097, 2019. Disponível em: <https://arxiv.org/pdf/1902.11097.pdf>. Acesso em 30 de agosto de 2020.

e suas formas mais comuns encontradas na prática. Em seguida, o texto trata da discriminação algorítmica propriamente dita.

#### 4.1 Viés algorítmico

Nos últimos anos surgiram evidências de que, no caso dos algoritmos aplicados à inteligência artificial, existe o risco (sobretudo não intencional) de ampliar condições discriminatórias<sup>64</sup> a partir do funcionamento “técnico” do algoritmo. O viés em algoritmos pode estar relacionado à sua autoria (a respeito do quão diversa é a equipe que elaborou o código de programação envolvido), pois a homogeneidade de perfil socioeconômico entre os programadores tende a reverberar em seu trabalho, do qual se originam algoritmos menos aptos a lidar com a diversidade de pessoas e situações que existem na prática. Também é uma causa relevante a maneira pela qual o algoritmo foi programado, pois é preciso que se leve em consideração o teor de equidade que o trabalho deve preservar. O propósito por trás da criação do algoritmo é igualmente relevante, pois embora suas diversas aplicações possam render enormes resultados financeiros aos seus criadores, é preciso observar que a inovação científica deve se comprometer com impactos positivos na sociedade, sem ampliar condições antijurídicas nem fortalecer retrocessos sociais.<sup>65</sup>

Conforme mencionado anteriormente, a ciência de dados<sup>66</sup> é a principal forma de alimentação e desenvolvimento das ferramentas atuais de inteligência artificial. Logo, não é surpreendente que muitos tipos de vieses causadores da discriminação algorítmica envolvam de uma maneira ou de outra o uso dos dados pelo sistema. O governo federal dos Estados Unidos, que há alguns anos vem elaborando estudos científicos para subsidiar sua política nacional de inteligência artificial, optou por agrupar os diferentes tipos de vieses em duas categorias principais<sup>67</sup>, que serão abordadas a seguir: a primeira relacionada aos dados de

---

<sup>64</sup> Os exemplos mais comuns de viés que os algoritmos apresentam em geral se relacionam a questões étnicas e raciais, ou de gênero e orientação sexual, como se demonstrará adiante.

<sup>65</sup> Essas são algumas das observações feitas por Joy Buolamwini, cientista da computação e pesquisadora vinculada ao Massachusetts Institute of Technology, nos Estados Unidos. Buolamwini fundou a Algorithmic Justice League, uma iniciativa para identificar e corrigir viés em algoritmos. Mais detalhes disponíveis em: <https://www.ajlunited.org/> Acesso em 9 de junho de 2020.

<sup>66</sup> A ciência de dados, por exemplo, permite utilizar algoritmos de aprendizado de máquina para possibilitar aos humanos identificar e compreender padrões em um conjunto de dados não estruturados (como textos, vídeos e imagens) que não haviam sido descobertos ou mesmo imaginados pelos próprios cientistas. Esta é, na verdade, uma das formas mais comuns da chamada análise preditiva, que tem provocado muito dos avanços recentes na área da inteligência artificial.

<sup>67</sup> THE WHITE HOUSE. *Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights*. Executive Office of the President, 2016. pp. 6 – 10. Disponível em:

entrada (*inputs*) em um algoritmo e a segunda abrangendo o funcionamento interno do algoritmo em si. É preciso observar, também, que existem diversas denominações para cada tipo de viés, mas como aparentemente (ainda) não há uniformidade ou padronização a respeito da sua nomenclatura, aqui será indicada somente sua conceituação, deixando de lado toda a diversidade de nomes que poderiam receber.

Em relação à primeira categoria, o viés se apresenta na medida em que o algoritmo precisa decidir quais dados serão utilizados pelo sistema, em detrimento de outros. Essa é uma ação corriqueira em todos os softwares que utilizam técnicas de inteligência artificial. Por exemplo, sua aplicação é comum em programas de navegação por satélite presentes em aparelhos (celulares) de telefonia móvel, nos quais os dados de velocidade de cada usuário permitem ao sistema conhecer a rota mais rápida para chegar ao destino escolhido. A situação descrita neste caso é benigna, mas sua análise permite compreender de modo simplificado as principais maneiras pelas quais o resultado da operação algorítmica pode resultar em práticas discriminatórias. Os vieses aqui podem ser causados por deficiências na inclusão e tratamento dos dados, isto é, quando algumas informações são consideradas como mais ou menos importantes para a decisão do algoritmo. No caso do exemplo apresentado, o sistema de navegação poderia não incluir a opção de realizar o trajeto por meio de bicicletas ou a pé, deixando em desvantagem as pessoas que não possuem carros.

Um outro tipo de viés muito frequente ocorre quando os dados apresentam algum erro ou estão desatualizados/incompletos. No caso do trajeto descrito, ainda que o sistema oferecesse a rota utilizando o transporte público, de nada adiantaria se as informações de horários não estivessem atualizadas, novamente prejudicando usuários que não utilizam carros.

Há ainda uma terceira categoria de viés algorítmico, relacionado aos dados de entrada, que compreende o viés de seleção (que talvez seja o mais frequente na prática), ocorrendo quando os dados inseridos no sistema não correspondem à realidade. No exemplo escolhido acima, se os dados forem majoritariamente obtidos a partir de telefones de última geração, é provável que a precisão do sistema seja prejudicada nas áreas mais pobres da região, que receberiam um serviço de menor qualidade, já que essa população de menor poder aquisitivo não participaria do serviço da mesma forma.

Por fim, um último tipo de viés que também é muito comum envolve a perpetuação de vieses históricos, por meio de *feedback loops*, isto é, nas situações em que um viés é replicado indefinidamente devido à capacidade da inteligência artificial de aprender com os padrões

detectados, conforme indicado no capítulo 2. Se o modelo de aprendizado está enviesado, os resultados estarão igualmente enviesados, replicando o formato equivocado que foi aprendido. Um exemplo comum são algoritmos de seleção de trabalhadores, que podem desconsiderar um candidato apto à uma vaga de emprego apenas por diferir da faixa etária normalmente contratada pela empresa que publicou a vaga.

Esses tipos de vieses devem ser considerados no desenvolvimento e aplicação de sistemas de inteligência artificial envolvidos na prestação de serviços, para que sejam preservados os critérios de eficiência, equidade e ética dispensados aos consumidores e/ou cidadãos em geral, em situações que envolvem tanto as instituições públicas quanto as privadas, como por exemplo concessão de crédito, licitações, alocação de recursos para programas de benefícios sociais e contratações de novos trabalhadores. Ainda não existem parâmetros oficiais nem consenso a respeito da melhor estratégia a ser adotada para mitigar os vieses em algoritmos. Contudo, medidas que promovam transparência e clareza de informação na escolha e tratamento dos dados são um ponto de partida importante para garantir que os dados inseridos em qualquer sistema algorítmico sejam precisos e adequados.

Prosseguindo a análise dos vieses algorítmicos em relação ao segundo ponto indicado anteriormente, que diz respeito ao funcionamento interno do algoritmo, é necessário lidar ainda com a dificuldade de que sua operação está distante da realidade de um usuário médio. Em regra, os códigos que compõem os sistemas algorítmicos são considerados confidenciais ou de propriedade das organizações que os desenvolveram e/ou naquelas onde são utilizados. Esta situação gera processos de tomada de decisão que apresentam diferentes graus de automação, mas que tem a característica comum de não apresentar transparência quanto à elaboração do resultado, o que deixa os usuários afetados com capacidade limitada para identificar os critérios pelos quais tais decisões foram tomadas. Logo, também se restringe suas opções de detectar e procurar corrigir quaisquer eventuais prejuízos ou erros gerados por algoritmos enviesados.

Esses vieses podem até mesmo significar que certos indivíduos serão excluídos inteiramente de determinadas oportunidades. Por exemplo, ao consultar páginas de busca na Internet essas pessoas podem (não) ter acesso a diferentes anúncios específicos de oportunidades profissionais, ou ofertas de mercadorias e serviços em geral, entretanto sem jamais suspeitar que essas oportunidades foram retiradas dos resultados de sua pesquisa a partir do funcionamento automático do algoritmo em uso pelo sistema de buscas. E se torna especialmente difícil para que os usuários tenham consciência do que está ocorrendo nesse tipo de situação, uma vez que se trata de um procedimento complexo, mas que pode se repetir frequentemente em outras interações dos usuários na rede. Novamente, os critérios de

transparência e clareza na informação devem nortear a atuação dos mecanismos de inteligência artificial. Sem medidas nesse sentido, as falhas causadas pelo funcionamento interno inadequado de sistemas de algoritmos se tornam mais difíceis de detectar e com maior probabilidade de se intensificar.

Uma forma de viés interno que os algoritmos podem apresentar está justamente relacionado aos sistemas de correspondência (*matching systems*)<sup>68</sup> mal-empregados na busca de informações, recursos ou serviços na rede. Por exemplo, mecanismos de pesquisa, plataformas de redes sociais e aplicativos em geral dependem de sistemas de correspondência para determinar os resultados da pesquisa realizada pelos seus usuários, além de decidir quais anúncios exibir e quais empresas recomendar. Tais sistemas de correspondência podem resultar em resultados discriminatórios se a estrutura (*design*) do sistema não for atualizada periodicamente ou ignorar vieses históricos e imprecisões nos dados ou algoritmos usados. Um exemplo disso envolve uma usuária francesa de um aplicativo de relacionamentos, que solicitou seus dados pessoais à empresa responsável e pode constatar como sua rede de relacionamentos afetivos estava sob “curadoria” do algoritmo em questão, que poderia desnecessariamente restringir as suas interações na plataforma.<sup>69</sup>

Outra forma de viés interno envolve os serviços de personalização e recomendação adotados em aplicativos e sistemas de busca que restringem ao invés de alargar as opções do usuário. Aqui, informações detalhadas sobre cada usuário individual são coletadas e analisadas para inferir suas preferências, interesses e crenças. A partir da categorização do usuário em um perfil específico, serão indicadas oportunidades, como novas músicas para baixar, vídeos para assistir, descontos no preço ou produtos para comprar com base no que o sistema entende que está de acordo com a preferência deste perfil. Estudos acadêmicos<sup>70</sup> têm demonstrado que os

---

<sup>68</sup> Existem diversos algoritmos diferentes que podem funcionar como um mecanismo de correspondência, dependendo do tipo de serviço prestado, como aplicativos de transportes ou de relacionamentos. Mas em geral esses sistemas realizam uma espécie de média ponderada que inclui enorme conjunto de dados diferentes, como histórico dos usuários, seu perfil de consumo, disponibilidade do prestador de serviço, etc. Por esta via o algoritmo consegue atingir a resposta que será recomendada a cada usuário de modo individual. Em regra, como acontece com todo sistema de *big data* e na estatística em geral, a precisão do resultado tende a aumentar junto com a quantidade e a qualidade dos dados ao qual se teve acesso. Mais detalhes disponíveis em: <https://www.entrepreneur.com/article/338442>. Acesso em 12 de novembro de 2020.

<sup>69</sup> A empresa alega que desde então alterou sua maneira de utilizar os dados pessoais de seus usuários, mas de forma geral se mantém no mercado esta lógica de funcionamento com base nos dados coletados a partir da atividade do usuário. Detalhes acerca do caso citado estão disponíveis no seguinte endereço eletrônico: <https://www.theguardian.com/technology/2017/sep/26/tinder-personal-data-dating-app-messages-hacked-sold>. Acesso em 22 de outubro de 2020.

<sup>70</sup> SWEENEY, L. *Discrimination in Online Ad Delivery*. Harvard University. January 28, 2013 <http://ssrn.com/abstract=2208240>; DATTA, A., TSCHANTZ, M. C., *Automated experiments on ad privacy settings: A tale of opacity, choice, and discrimination*. In Proceedings on Privacy Enhancing Technologies (PoPETs). 2015. <https://www.andrew.cmu.edu/user/danupam/dtd-pets15.pdf> Acesso em 19 de outubro de 2020.

algoritmos de recomendação deste conteúdo podem inadvertidamente restringir a disponibilidade de informações a determinados grupos sociais, deixando-os sem as mesmas oportunidades de acesso econômico e inclusão em comparação aos demais. Um exemplo é o acesso ao crédito, onde os consumidores que não possuem histórico de contratação de produtos financeiros considerados suficientes pelo algoritmo terão maior probabilidade de ter uma oferta recusada em virtude do seu padrão de consumo. Como se poderia imaginar, esse perfil de consumidor é composto principalmente pelo segmento da população menos privilegiado socioeconomicamente.<sup>71</sup>

A terceira categoria de viés interno abrange os mecanismos de tomada de decisão que pressupõem equivocadamente que a correlação estatística necessariamente implica causalidade<sup>72</sup>. Isto é, um programador (humano) ou o próprio sistema algorítmico pode analisar dois fatores que frequentemente ocorrem juntos (por exemplo, pertencer a determinado grupo étnico e ter um certo nível de renda), e presumir que há necessariamente algum nexo de causalidade entre os dois. Logo, criar uma relação causal quando há somente correlação entre os dados pode levar à discriminação de usuários específicos. Um exemplo ocorre em análises demográficas promovidas por seguros de saúde feita nos Estados Unidos: para uma mesma classificação de risco de adoecer, os pacientes negros na verdade já apresentam condições de saúde mais agravadas do que os pacientes brancos. Essa diferença ocorre porque os negros em geral possuem uma renda menor e, portanto, tem menos acesso aos serviços de saúde (privados) naquele país.<sup>73</sup> Uma das consequências é que os médicos podem utilizar essa classificação de risco enviesada para formular o tratamento mais adequada para cada paciente, o que no caso dos negros pode estar aquém da sua real necessidade. Embora o estudo seja estrangeiro,

---

<sup>71</sup> THE WHITE HOUSE. Idem, ibidem. p 11. Disponível em:

[https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/2016\\_0504\\_data\\_discrimination.pdf](https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf). Acesso em 30 de agosto de 2020.

<sup>72</sup> A estatística inferencial faz uso da análise descritiva dos dados de uma amostra para chegar a conclusões sobre um grupo maior, ao qual não se teve acesso. O objetivo é determinar a probabilidade com que o resultado da análise da amostra se aplica a toda a universalidade dos dados disponíveis. Isto é, criar um modelo que permite prever tendências e comportamentos dos dados que compõem aquele conjunto determinado em estudo. A relação de causalidade pode ser demonstrada, por exemplo, por meio de experimentos aleatórios e controlados. Este é um assunto complexo e que envolve frequentemente grande risco de erro, mas em regra as análises demonstram que duas variáveis associadas nem sempre apresentam relação de causa e efeito entre si, e que quaisquer conclusões a esse respeito devem ser feitas com cautela. Para mais detalhes sobre análise de causalidade em estatística consultar: PEARL, Judea. *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press, 2000. Disponível em: <http://bayes.cs.ucla.edu/BOOK-99/book-toc.html>. Acesso em 19 de outubro de 2020.

<sup>73</sup> OBERMEYER, Ziad; POWERS, Brian; VOGELI, Christine; MULLAINATHAN, Sendhil. *Dissecting racial bias in an algorithm used to manage the health of populations*. Science n° 366, 2019. pp. 447-453. Disponível em:

<https://www.ehdc.org/sites/default/files/resources/files/Dissecting%20racial%20bias%20in%20an%20algorithm%20used%20to%20manage%20the%20health%20of%20populations.pdf>. Acesso em 22 de outubro de 2020.

considerando as condições socioeconômicas do Brasil provavelmente as mesmas conclusões podem ser aplicadas por aqui.

Finalmente, há também uma quarta categoria de viés interno, que ocorre quando o sistema trabalha com conjuntos de dados que não correspondem à realidade daquilo que está sendo representado, isto é, quando as informações estão incompletas ou representam desproporcionalmente certos estratos da população. Neste cenário, o resultado da operação algorítmica será distorcido, sedimentando de modo eficaz a discriminação devido à falha dos dados de entrada iniciais. Considerando que a disponibilidade de dados, o acesso à tecnologia e a participação nas ferramentas digitais estão distribuídos de maneira desigual na sociedade, é necessário que os algoritmos sejam elaborados de modo a lidar com essa diversidade. Do contrário, barreiras econômicas, linguísticas, estruturais ou socioeconômicas podem ser ampliadas pelos mecanismos de inteligência artificial. Sem ajuste adequado, essa característica pode se tornar uma falha sistêmica que reforça os padrões existentes de discriminação por meio de uma descrição estatística pouco representativa que será utilizada pelo algoritmo, gerando tratamento desigual para as pessoas envolvidas. No Reino Unido, por exemplo, o algoritmo utilizado na seleção de candidatos ao visto de imigração teve seu uso descontinuado, no que se tornou talvez o primeiro desafio bem-sucedido a um sistema de tomada de decisão de inteligência artificial naquele país. As críticas se concentraram no fato de que o software estaria reproduzindo as práticas racistas que foram institucionalizadas em parte do século passado quanto à origem nacional dos imigrantes.<sup>74</sup> Um episódio em particular, conhecido como *Windrush scandal* (escândalo de Windrush), foi a grande motivação que levou o governo britânico a alterar sua política migratória.<sup>75</sup>

Esses tipos de vieses podem representar grandes problemas para os usuários, considerando a disseminação cada vez maior das técnicas de inteligência artificial (como as apresentadas no capítulo 2). No caso do aprendizado de máquina, por exemplo, os códigos de programação são complexos e por vezes incompreensíveis até mesmo para seus programadores. Não obstante, essa ferramenta é cada vez mais usada em situações do cotidiano, como análise de crédito, empréstimos e contratação de novos trabalhadores, conforme indicado

---

<sup>74</sup> Mais detalhes disponíveis em: <https://www.theguardian.com/uk-news/2020/aug/04/home-office-to-scrap-racist-algorithm-for-uk-visa-applicants>. Acesso em 24 de outubro de 2020.

<sup>75</sup> Neste caso houve a prisão e deportação ilegal de cidadãos britânicos negros, devido à interpretação equivocada da legislação vigente nas décadas de 1960 e 1970, que apresentava critérios hoje considerados racistas. Mais detalhes sobre este episódio disponíveis na avaliação independente que foi promovida pelo governo, disponível no seguinte endereço eletrônico: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/876336/6.5577\\_HO\\_Windrush\\_Lessons\\_Learned\\_Review\\_LoResFinal.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/876336/6.5577_HO_Windrush_Lessons_Learned_Review_LoResFinal.pdf).

anteriormente. Como esses algoritmos se tornam cada vez mais sofisticados, também fica mais difícil preservar a transparência e explicar com clareza o método pelo qual as máquinas tomam decisões autônomas. Assim, programadores e cientistas de dados podem inadvertidamente ou inconscientemente projetar, treinar ou implantar sistemas de *big data* enviesados, o que é uma consequência inevitável de sua própria cultura e histórico de vida, como acontece com qualquer ser humano.

Os vieses dos algoritmos estão muito associados aos vieses cognitivos que os seres humanos apresentam frequentemente em seu próprio raciocínio, e que são transferidos ao código de programação pelo trabalho do programador ou pelos dados que alimentarão o funcionamento do algoritmo. Os vieses cognitivos atuam como atalhos ou desvios no processo racional do pensamento humano, consistindo em um caminho “mais curto” do qual o cérebro se vale para simplificar a realidade e resolver determinado problema com menos esforço, atingindo a resposta mais rápida, que nem sempre é a correta. Em relação à inteligência artificial, isso poderia ser mitigado pela transparência do código de programação, o que dificilmente ocorre por razões de propriedade intelectual.<sup>76</sup> Por isso a mitigação de vieses deve ser uma preocupação desde a fase de projeto do software, para os usuários se sintam seguros e confiantes, sabendo que sistemas algorítmicos enviesados são a exceção, e não a regra geral.

Deste modo, já que o uso da inteligência artificial deve se tornar cada vez mais presente na vida cotidiana das pessoas, será necessário dedicar mais atenção para compreender a maneira de operação desses sistemas, avaliando continuamente os dados que alimentam o algoritmo e seu mecanismo de atuação, bem como os resultados produzidos. Inclusive, ainda é comum a crença de que os números seriam “naturalmente” neutros e imparciais, e que representariam sempre a mesma realidade objetiva, o que traz sérias dificuldades e contribui para ofuscar situações discriminatórias que afetam negativamente a vida das pessoas.<sup>77</sup> Assim, os desafios relacionados a essa nova realidade tecnológica incluem sobretudo os resultados das operações automatizadas, que devem sempre garantir que as informações sobre lugares, pessoas,

---

<sup>76</sup> Sobre este ponto, consultar: STEFFENS, Luana. *A influência dos vieses cognitivos na tomada de decisão pela inteligência artificial: um estudo baseado nas evidências no caso norte-americano COMPAS*. In: Fabio da Silva Veiga; Denise Pires Fincato. (Org.). Estudos de Direito, Desenvolvimento e novas Tecnologias. 1ed. Porto Alegre: Instituto Iberoamericano de Estudos Jurídicos, 2020, v. 1, p. 1-419.

<sup>77</sup> Para o antropólogo Rodrigo Ochigame, o pressuposto de que os números são inerentemente neutros e imparciais serve para ocultar determinadas lutas políticas e reduzir a compreensão do fenômeno da discriminação algorítmica a um critério puramente matemático. Assim, após ajustes no modelo do código de programação utilizado, o algoritmo poderia ser apresentado como uma solução acabada para conflitos de caráter político, como discriminação de minorias. A análise completa de Ochigame pode ser consultada no seguinte endereço eletrônico: <https://phenomenalworld.org/analysis/long-history-algorithmic-fairness>. Acesso em 25 de outubro de 2020.

comportamentos, etc. sejam usados de forma adequada, ética e na promoção de princípios democráticos, como a não-discriminação e igualdade de oportunidades.

## 4.2 Discriminação algorítmica

É claro que as tecnologias que fazem uso da ciência de dados também podem apresentar o potencial de aprimorar a capacidade de detecção e prevenção de discriminação em geral. Como visto acima, existem perspectivas que inclusive defendem o uso de algoritmos para ajudar os seres humanos em seu processo de tomada de decisão, com o objetivo de preservar critérios de equidade e objetividade<sup>78</sup>, ou que até mesmo não recomendam a possibilidade da revisão humana de decisões automatizadas, o que será abordado no próximo capítulo. Porém, o problema abordado nesta pesquisa diz respeito aos algoritmos que não funcionam tão bem na prática e podem ampliar situações discriminatórias numa escala maior na medida que em os mecanismos de inteligência artificial são adotados pela sociedade em diversas situações.

Assim, quando essas aplicações tecnológicas não são implementadas de modo adequado, a inteligência artificial pode contribuir na prática para perpetuar, exacerbar ou mascarar situações discriminatórias, conforme indicado acima. Nos últimos anos a questão do viés na operação algorítmica vem assumindo uma posição de destaque nas pesquisas acadêmicas e debates legislativos que envolvem o *big data*. Quando se utiliza a ciência de dados para tratar vastas quantidades de dados (alguns dos quais podem nunca se tornar ‘informação’ de fato), sem que exista interação humana no processo de tomada de decisão do algoritmo, é preciso se considerar qual será o impacto causado.<sup>79</sup>

Atualmente, há uma percepção entre os pesquisadores<sup>80</sup> de que algoritmos podem potencializar diversas espécies de discriminação e promover violações de direitos

---

<sup>78</sup> “There is a large body of research on algorithmic decision making that dates back several decades. And the existing studies on this topic all have a remarkably similar conclusion: Algorithms are less biased and more accurate than the humans they are replacing.” O autor Alex P. Miller reconhece que algoritmos podem ser parte da solução, embora seja sempre necessário preservar a transparência e a ética na elaboração do código e no uso dos dados. Artigo disponível em: <https://hbr.org/2018/07/want-less-biased-decisions-use-algorithms>. Acesso em 25 de outubro de 2020.

<sup>79</sup> CARRUTHERS, Caroline, JACKSON, Peter.; *The Chief Data Officer’s Playbook*. EBSCO Publishing : eBook Collection, 2018.

<sup>80</sup> Esta é a perspectiva de autores abordados ao longo deste capítulo 4, como Cathy O’Neil, Meredith Broussard, Safiya Noble, Tarcízio Silva, dentre outros.

fundamentais. Existem exemplos<sup>81</sup> na área penal, eleitoral<sup>82</sup>, consumerista, etc. que justificam amparo jurídico para sanar as violações originadas pela operação enviesada dos algoritmos. O problema se torna mais sensível na medida em que esses algoritmos são quase sempre desenvolvidos por empresas privadas e que, portanto, detêm a sua propriedade<sup>83</sup>, de modo que seu mecanismo de funcionamento nem sempre pode ser auditado, como já mencionado.<sup>84</sup> A Tabela 1 a seguir apresenta alguns exemplos de vieses raciais causados por algoritmos.<sup>85</sup>

Tabela 1 - Exemplos concretos de racismo algorítmico categorizados pelo conteúdo específico de discriminação racial que apresentam

<b>Caso de racismo algorítmico</b>	<b>Conteúdo da discriminação</b>
Sistema do Google permite empresas exibirem anúncios sobre crime especificamente a afro-americanos	Suposição de criminalidade
Resultados no Google Imagens apresentam hiper-sexualização para buscas como “garotas negras”	Exotização; negação de

<sup>81</sup> Esses exemplos foram identificados no noticiário eletrônico (brasileiro e estrangeiro), disponível em: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>; <https://www.nexojournal.com.br/externo/2019/11/24/A-parcialidade-dos-algoritmos>; <https://www.theatlantic.com/technology/archive/2011/04/the-curious-connection-between-apps-for-gay-men-and-sex-offenders/237340>; Acesso em 25 de outubro de 2020.

<sup>82</sup> O Provimento nº 71, de 13 de junho de 2018, editado pela Corregedoria Nacional de Justiça inclusive vedou manifestações de magistrados favoráveis ou contrárias a candidatos e partidos políticos que possam ser entendidas como exercício de atividade político-partidária. Esse entendimento foi confirmado pelo Ministro do Supremo Tribunal Federal Roberto Barroso, afirmando que “tais declarações em redes sociais, com a possibilidade de reprodução indeterminada de seu conteúdo e a formação de algoritmos de preferências, contribuem para se alcançar um resultado eleitoral específico, o que é expressamente vedado pela Constituição.” (Medida Cautelar em Mandado de Segurança 35793/Distrito Federal, Relator Ministro Roberto Barroso, julgado em 04/09/2018, publicado 06/09/2018).

<sup>83</sup> O Artigo 5º, incisos XXVII, XXVIII e XXIX da Constituição de 1988 juntamente com o Artigo 7º da Lei 9.610 de 1998 (Lei de Direitos Autorais) e a Lei 9.609 de 1998 (Lei do Software) resguardam ao criador o direito de propriedade sobre o código de programação de um algoritmo.

<sup>84</sup> “Se o policiamento preditivo significa que alguns indivíduos terão mais envolvimento policial em suas vidas, é preciso haver um mínimo de transparência”, disse em entrevista Adam Schwartz, advogado sênior da Electronic Frontier Foundation. Até que façam isso, o público não deve confiar que os dados inseridos e os algoritmos são uma base sólida para prever qualquer coisa.” Tradução livre de “If predictive policing means some individuals are going to have more police involvement in their life, there needs to be a minimum of transparency,” Adam Schwartz, a senior staff attorney with the Electronic Frontier Foundation, said in an interview “Until they do that, the public should have no confidence that the inputs and algorithms are a sound basis to predict anything.” Disponível em: <https://www.mic.com/articles/156286/crime-prediction-tool-pred-pol-only-amplifies-racially-biased-policing-study-shows>. Acesso em 25 de outubro de 2020.

<sup>85</sup> A tabela foi retirada (com adaptações) a partir do seguinte artigo: SILVA, Tarcízio. *Racismo Algorítmico em Plataformas Digitais: microagressões e discriminação em código*. In: Anais do IV Simpósio Internacional LAVITS – Assimetrias e (In)visibilidades: Vigilância, Gênero e Raça. Salvador, Bahia, Brasil, 2019. Disponível em: [https://www.researchgate.net/publication/333700308\\_Racismo\\_Algoritmico\\_em\\_Plataformas\\_Digitais\\_microagressoes\\_e\\_discriminacao\\_em\\_codigo](https://www.researchgate.net/publication/333700308_Racismo_Algoritmico_em_Plataformas_Digitais_microagressoes_e_discriminacao_em_codigo). Acesso em 25 de outubro de 2020.

	cidadania
Facebook esconde manifestações contra violência policial racista	Negação de realidades raciais
Google Photos marca fotos de jovens negros com a tag “gorila”	Negação de cidadania
Chatbot da Microsoft torna-se racista em menos de um dia	Diversas
Robôs conversacionais de startups não encontram face de mulher negra; sistemas de visão computacional erram gênero e idade de mulheres negras	Negação de cidadania; exclusão e isolamento
Mecanismos de busca de bancos de imagens invisibilizam famílias e pessoas negras	Negação de realidades raciais
App que transforma selfies equipara beleza à brancura	Exotização; exclusão e isolamento
APIs de visão computacional confundem cabelo negro com perucas	Exotização
Ferramentas de processamento de linguagem natural possuem vieses contra linguagem e temas negros	Patologização de valores culturais
Análise facial de emoções associa categorias negativas a atletas negros	Suposição de criminalidade
Twitter decide não banir discurso de ódio nazista/supremacista branco para não afetar políticos republicanos nos Estados Unidos	Negação de realidades raciais; exclusão

Fonte: Silva, 2019.

Neste contexto, a sociologia é uma das principais áreas científicas que tem estudado o fenômeno da discriminação, e o trabalho clássico do sociólogo alemão Norbert Elias em uma pequena cidade no interior da Inglaterra contribuiu para estabelecer o conceito de *outsider*.<sup>86</sup>

<sup>86</sup> Em sua pesquisa de campo, Norbert Elias explorou conceitos como superioridade social e moral, autopercepção e reconhecimento, além de pertencimento e exclusão em um grupo social que à primeira vista era bastante homogêneo. ELIAS, Norbert. *Os estabelecidos e os outsiders: sociologia das relações de poder a partir de uma pequena comunidade*. Rio de Janeiro: Jorge Zahar, 2000.

Sua pesquisa facilitou a compreensão das distinções de valor que os indivíduos acabam por atribuir entre si dentro de um determinado grupo social, em uma espécie de “hierarquia classificatória” (*status quo*) que apresenta profundas implicações na vida de cada pessoa que compõe esse mesmo grupo. Para Elias, essa estrutura corresponderia a um modelo geral que explicaria (ao menos em parte) a motivação para atitudes discriminatórias dos seres humanos em relação aos indivíduos percebidos como integrantes de um estrato social inferior. Este fato também serve para entender as crescentes tensões sociais e políticas existentes no atual contexto de globalização, pois as pessoas em regra não foram preparadas para a intensificação dos processos de mobilidade social, migrações, mudanças tecnológicas, etc. que vem ocorrendo em todo o mundo.

Mais especificamente sobre a questão envolvendo a discriminação algorítmica, a cientista social norte-americana Safiya Umoja Noble<sup>87</sup> aponta para a falta de consciência por parte da sociedade em geral em relação à atuação silenciosa dos algoritmos enviesados que podem produzir diversos casos de discriminação, como visto acima. Segundo a autora, a atuação discriminatória desses mecanismos de inteligência artificial pode se tornar umas das questões mais relevantes envolvendo os direitos humanos no século XXI.

Embora Noble tenha voltado sua atenção até o momento para o fator humano por detrás da engenharia de software, o principal risco de viés aqui pode ser considerado não-intencional e difícil de detectar, como já demonstrado anteriormente. Ou seja, a análise da autora privilegia a perspectiva de que os vieses inerentes aos programadores humanos (que elaboram os códigos de programação para compor os algoritmos de inteligência artificial) representam uma das principais causas da discriminação algorítmica. Neste sentido, talvez a principal responsabilidade desses profissionais seja a de garantir que esses vieses sejam corrigidos ao longo do desenvolvimento do programa, uma vez que é praticamente impossível treinar seres humanos para não apresentar nenhum tipo de viés. Além disso, ainda que o algoritmo funcione bem, isso não impede que os dados estejam enviesados e comprometam a qualidade da operação ao final, como indicado anteriormente.

De todo modo, Noble também insiste na transparência algorítmica enquanto um meio de fomentar o debate público a respeito de como lidar com os impactos sociais, políticos e econômicos da utilização de mecanismos de inteligência artificial da maneira mais eficaz, por meio de políticas públicas específicas para este tema. A autora também tem o mérito de chamar

---

<sup>87</sup> NOBLE, Safiya Umoja. *Algorithms of oppression: how search engines reinforce racism*. New York: New York University Press, 2018.

atenção para o fato de que a tecnologia não deve contribuir para naturalizar a discriminação a qual determinados grupos sociais (sobretudo minorias raciais) estão submetidos historicamente.

Essa perspectiva é compartilhada por cientistas de dados como Catherine (“Cathy”) O’Neil, que desenvolveu um estudo sobre falhas em modelos matemáticos utilizados pelo *big data* para compreender (e antecipar) padrões de comportamento humano por meio de algoritmos. Em sua carreira profissional, O’Neil adquiriu progressivamente a percepção de que muitas suposições equivocadas estão disfarçadas nas fórmulas matemáticas e terminam compondo o código de programação sem teste ou reflexão prévia por parte dos profissionais envolvidos em sua elaboração. Na visão da autora, isso ocorre devido ao fato de que os eventuais prejudicados (pela decisão automatizada enviesada) apresentam dificuldade de comprovar suas suspeitas ao mesmo tempo em que existe certa presunção de tecnicidade e objetividade em relação aos algoritmos. A principal razão para este desinteresse das empresas estaria associada à perspectiva de que os lucros gerados pelo negócio desestimulam a busca pela revisão ou investigação de eventuais problemas envolvendo a ferramenta de inteligência artificial utilizada.

Neste caso, um dos tipos mais comuns de viés apontados por O’Neil envolve o *feedback loop*, conforme descrito nas páginas anteriores. Sem contraponto com ao menos parte da realidade, não há como saber se as informações geradas pelo mecanismo estatístico do algoritmo estão na verdade produzindo análises defeituosas e prejudiciais, sem nunca de fato aprender com seus erros. A autora aponta que na prática alguns mecanismos de inteligência artificial acabam por criar uma realidade paralela que será usada para justificar (matematicamente) os seus próprios resultados. Esse tipo de modelo tende a se retroalimentar e se perpetuar enquanto não houver correção dos dados. Nas palavras de O’Neil:<sup>88</sup>

Como deuses, esses modelos matemáticos eram opacos, seu funcionamento invisível para todos, exceto para os sacerdotes mais elevados em seu domínio: matemáticos e cientistas da computação. Seus veredictos, mesmo quando errados ou prejudiciais, estavam fora de discussão ou apelação. E eles tendiam a punir os pobres e oprimidos em nossa sociedade, enquanto tornavam os ricos mais ricos. (...) Isso ocorre, em parte, porque eles são projetados para avaliar um grande número de pessoas. Eles se

---

<sup>88</sup> “Like gods, these mathematical models were opaque, their workings invisible to all but the highest priests in their domain: mathematicians and computer scientists. Their verdicts, even when wrong or harmful, were beyond dispute or appeal. And they tended to punish the poor and the oppressed in our society, while making the rich richer. (...) This is, in part, because they are engineered to evaluate large numbers of people. They specialize in bulk, and they’re cheap. That’s part of their appeal. The wealthy, by contrast, often benefit from personal input. A white-shoe law firm or an exclusive prep school will lean far more on recommendations and face-to-face interviews than will a fast-food chain or a cash-strapped urban school district. The privileged, we’ll see time and again, are processed more by people, the masses by machines.” O’NEIL, Cathy. *Weapons of math destruction: how big data increases inequality and threatens democracy*. New York: Crown, 2016. pp 10-14.

especializam em grandes quantidades e são baratos. Isso é parte de seu apelo. Os ricos, em contraste, frequentemente se beneficiam de análises individualizadas. Um escritório de advocacia ou uma escola preparatória exclusiva se apoiarão muito mais em recomendações e entrevistas cara a cara do que uma rede de fast food ou uma rede escolar urbana sem dinheiro. Os privilegiados, veremos repetidamente, recebem tratamento majoritariamente feito por pessoas, as massas por máquinas. (tradução livre)

Ainda, jornalistas como Meredith Broussard<sup>89</sup> chamam a atenção para todas as consequências inesperadas que novas tecnologias trazem consigo, ao relatar um caso onde um *drone* foi alvejado ao sobrevoar a casa de uma pessoa nos Estados Unidos. O dono da casa se sentiu ameaçado e, desconhecendo a origem e propriedade do objeto, atirou no robô aéreo enquanto este sobrevoava sua residência. A polícia foi chamada até o local, mas não soube lidar com o fato, e os agentes decidiram prender o proprietário do imóvel por ser quem estava armado. Por fim, um tribunal posteriormente o inocentou por considerar que exercia seu direito legítimo de defender sua propriedade. Este fato é utilizado pela autora para evidenciar que quase sempre as empresas que introduzem novas tecnologias no mercado não estão conscientes dos riscos envolvidos na utilização de seus produtos, mesmo diante do fato bastante conhecido de que a população norte-americana é uma das mais armadas do mundo.<sup>90</sup>

No caso da discriminação algorítmica, menos inteligível ao público leigo do que o exemplo relatado acima, Broussard afirma que muitos profissionais e empresas possuem uma visão ingênua sobre como as novas tecnologias serão usadas. Sem a cautela necessária, a autora acredita que as novas ferramentas de inteligência artificial podem prestar um serviço incompatível com o interesse público, seguindo o raciocínio exposto nos parágrafos anteriores.

Como se sabe, o ordenamento jurídico brasileiro é refratário a quaisquer formas de preconceito e discriminação, seguindo as determinações constitucionais que promovem a igualdade entre os cidadãos. O que não significa, claro, que toda forma de discriminação seja negativa ou inválida, pois tanto a igualdade formal (isonomia em sentido estrito) quanto a igualdade material (que está associada à ideia de justiça distributiva e social) devem ser atendidas pelo Estado.

Na visão de Luís Roberto Barroso,<sup>91</sup>

---

<sup>89</sup> BROUSSARD, Meredith. *Artificial unintelligence: how computers misunderstand the world*. Cambridge: MIT Press, 2018.

<sup>90</sup> BROUSSARD, Meredith. *Idem*, *ibidem*. p. 77 – 78.

<sup>91</sup> BARROSO, Luís Roberto. *Diferentes, mas iguais: o reconhecimento jurídico das relações homoafetivas no Brasil*. Revista Brasileira de Direito Constitucional, v. 17, p. 105-138, 2011. Disponível em: [http://www.luisrobertobarroso.com.br/wp-content/uploads/2017/09/diferentes\\_mas\\_iguais\\_atualizacao\\_2011.pdf](http://www.luisrobertobarroso.com.br/wp-content/uploads/2017/09/diferentes_mas_iguais_atualizacao_2011.pdf). Acesso em 26 de outubro de 2020.

a Constituição Federal de 1988 consagra o princípio da igualdade e condena de forma expressa todas as formas de preconceito e discriminação. A menção a tais valores vem desde o preâmbulo da Carta, que enuncia o propósito de se constituir uma “sociedade fraterna, pluralista e sem preconceitos”. O art. 3º renova a intenção e lhe confere inquestionável normatividade, enunciando serem objetivos fundamentais da República “construir uma sociedade livre, justa e solidária” e “promover o bem de todos, sem preconceitos de origem, raça, sexo, cor, idade e quaisquer outras formas de discriminação”. O caput do art. 5º reafirma que “todos são iguais perante a lei, sem distinção de qualquer natureza”. O constituinte incluiu, ainda, menções expressas de rejeição ao racismo e à discriminação contra as mulheres.

Neste ponto, o objeto desta pesquisa trata da questão da igualdade material, que envolve enormes dificuldades teóricas e práticas na sua concretização. Neste sentido, Barroso defende que a concreção do direito geral de igualdade envolve três dimensões principais: formal, material e também a igualdade enquanto reconhecimento (dimensão simbólica que impacta na autoestima dos indivíduos).<sup>92</sup> Essas três grandes categorias precisam ser atendidas para que a desigualdade seja reduzida em um país como o Brasil, marcado historicamente por profundas diferenças entre seus cidadãos. De todo modo, na visão de Barroso é necessário haver intervenção estatal para garantir a efetividade dos mandamentos constitucionais em âmbito público e privado, pois<sup>93</sup>

onde não exista um motivo relevante e legítimo que justifique diferença no tratamento, a equiparação deve ser a conduta de todos os órgãos e agentes públicos e, dentro de certa medida, deve ser imposta até mesmo aos particulares.

Portanto, promover a igualdade material requer políticas públicas de caráter universalista, conduzidas pelo Estado, como é o caso das ações afirmativas, por exemplo. A atuação do poder público é essencial para mitigar situações históricas de discriminação contra determinados grupos sociais. Além disso, a garantia da igualdade em todas as suas dimensões é algo que compõe o próprio conceito de democracia, ao assegurar igualdade de oportunidades. Ela também possibilita o livre desenvolvimento da personalidade de todos os cidadãos e favorece o crescimento do indivíduo, inserido em uma sociedade plural e diversificada.

Assim, em síntese, este capítulo abordou alguns dos principais tipos de vieses que originam situações discriminatórias em mecanismos de inteligência artificial. Prosseguindo com o raciocínio exposto acima, a análise agora se volta para a regulação da inteligência

---

<sup>92</sup> Conforme voto proferido no julgamento da Ação Declaratória de Constitucionalidade 41, julgada pelo Plenário do STF em 08 de junho de 2017.

<sup>93</sup> BARROSO, Luís Roberto. Idem, *ibidem*.

artificial enquanto possível estratégia para oferecer condições que possam neutralizar a discriminação algorítmica.

## 5 Possíveis soluções regulatórias para o Brasil

Até este ponto, foi apresentado de modo breve o funcionamento de algumas das principais ferramentas de inteligência artificial e dos eventuais vieses que podem apresentar. Neste capítulo final o objetivo específico consiste em investigar se os mecanismos jurídicos disponíveis são suficientes para lidar com os problemas associados ao viés algorítmico.

Conforme mencionado ao longo do texto, a Constituição atual tem como um de seus objetivos fundamentais garantir a não-discriminação entre os indivíduos no Brasil. Porém, como se sabe, a mera disposição da Lei Maior nem sempre é suficiente para garantir sua eficácia. Assim, deste ponto em diante o texto busca identificar o método mais eficaz de concretizar o mandamento citado. Considerando que o Brasil não dispõe de uma política pública especificamente voltada para o tema do viés algorítmico, surge a preocupação de como conferir eficácia ao texto constitucional e garantir a não-discriminação no uso da inteligência artificial. Ao longo das próximas páginas, a abordagem desta pesquisa se volta especialmente para a questão da revisão humana de decisões automatizadas, entendida aqui como um método possível de mitigação dos prejuízos causados pelos algoritmos enviesados e que ficou de fora da redação final da LGPD, conforme já indicado no capítulo 2.

Neste contexto, o argumento principal deste trabalho diz respeito à necessidade de uma política pública específica que esteja apta a concretizar a respectiva disposição constitucional de garantia da não-discriminação. Isto requer uma atuação positiva por parte do legislador para garantir a efetividade deste mandamento, por se tratar de norma que não possui eficácia direta e imediata, como já indicado.<sup>94</sup>

Como visto anteriormente, o Marco Civil da Internet não tem incidência específica sobre a questão da discriminação algorítmica. A LGPD, por sua vez, deixou de oferecer a obrigatoriedade de revisão humana de decisões automatizadas como um recurso que poderia ser de grande valia para remediar administrativamente eventuais prejuízos aos indivíduos afetados. Deste modo, este capítulo analisa inicialmente a capacidade do próprio setor privado de se adequar voluntariamente ao arcabouço previsto pela Constituição, sobretudo por meio das políticas internas de ética das grandes empresas de tecnologia. Após, a atenção é dedicada ao sistema regulatório vigente propriamente dito, para identificar eventuais insuficiências normativas. É desnecessário insistir, entretanto, que os direitos fundamentais previstos no

---

<sup>94</sup> Sobre este assunto, ver a nota de número 62.

ordenamento jurídico brasileiro permanecem como o principal meio para proteger as pessoas de situações discriminatórias.<sup>95</sup>

### 5.1 A dimensão ética

No campo filosófico, a ética tem difícil definição, embora esteja presente em todas as facetas da vida humana. Isto ocorre porque o ser humano possui capacidade de reflexão sobre os seus próprios atos, o que lhe confere a possibilidade de escolha, de modo consciente. Essa liberdade de agência é um dos fundamentos da ética na medida em que trabalha com a operação sobre alternativas, ou seja, qualquer tomada de decisão requer um grau mínimo de consciência da opção escolhida. A faculdade de escolher, então, consiste em uma obrigação imposta ao indivíduo, já que o processo decisório em geral apresenta uma multiplicidade de caminhos onde é necessário avaliar e decidir por qual seguir.

Deste modo, a tomada de decisão humana supõe a avaliação dos múltiplos fatores que perfazem uma situação complexa, perpassada por valores morais, estéticos, religiosos, de utilidade, dentre outros. Na visão de Eduardo Bittar, atualmente as relações econômicas e de reprodução social da vida material humana constituem uma das principais determinantes das possibilidades éticas. O autor argumenta que “uma palavra em desuso como esta – ética –, apesar de seu valor, tem sido considerada simplesmente um obstáculo a mais a atravancar o andamento das facilidades do pragmatismo consumista e comercial.”<sup>96</sup>

Como acontece em todos os segmentos econômicos e ramos profissionais, a preocupação com princípios éticos a respeito do trabalho desenvolvido é essencial para garantir resultados compatíveis com o interesse público. Em relação ao *big data*, a ética desempenha um papel fundamental para assegurar que sua utilização esteja de acordo com valores positivos para a sociedade, uma vez que a automação do processo de análise e tomada de decisões pode trazer amplas consequências para as pessoas envolvidas, conforme demonstrado no capítulo anterior. A responsabilidade e o compromisso ético da ciência de dados nesse ponto é ainda maior:<sup>97</sup>

---

<sup>95</sup> Sobre este assunto, ver a nota de número 45.

<sup>96</sup> BITTAR, Eduardo Carlos Bianca. *Curso de ética jurídica: ética geral e profissional*. 13ª ed. São Paulo: Saraiva, 2016. p. 18.

<sup>97</sup> “Making ethical mistakes can have a big impact on the future of how we treat data. If we get it wrong public perception can change really quickly and have a very negative impact on what we can do; laws can change and enforce compliance but as the use of data evolves at such a pace, will the laws that legislate it cover what the data needs them to cover? In some ways we are already trying to deal with this at the legislative level, but are we addressing the core problems we are trying to fix or are we constraining ourselves too much? And is this the price we are paying for not taking this area and its security seriously enough? Yet we have to progress, that is

Cometer erros éticos pode ter um grande impacto no futuro de como tratamos os dados. Se errarmos, a percepção do público pode mudar muito rapidamente e ter um impacto muito negativo sobre o que podemos fazer; as leis podem mudar e exigir conformidade, mas à medida que o uso de dados evolui tão rapidamente, as leis que o regem terão abrangência sobre o que os dados precisam que elas cubram? De certa forma, já estamos tentando lidar com isso no nível legislativo, mas estamos abordando os problemas centrais que tentamos resolver ou estamos nos restringindo demais? E é esse o preço que pagamos por não levar esta área e sua segurança a sério o suficiente? No entanto, temos que progredir, isso é inevitável; a estagnação não serve para nada, então como podemos inovar com responsabilidade? (...) O primeiro passo é entender o papel que desempenhamos em tudo isso e assumir a responsabilidade por isso. Pense nas consequências tanto quanto possível, tanto as diretas quanto as não intencionais, e monitore a reação a elas. Procurar coisas desconhecidas é uma posição realmente difícil de se estar, mas se você não olhar, provavelmente nunca as verá. Pelo menos monitorar o que está acontecendo lhe dá essa chance de lutar. Em última análise, você tem que agir. Este é um campo em evolução, não tenha medo de admitir se as coisas não saírem do jeito que você queria e precisar corrigir a trajetória. (tradução livre)

De certa forma, é perceptível que as considerações éticas venham se tornando uma preocupação cada vez mais relevante na atuação das empresas em geral, e não apenas por razões de imagem para conquistar boa recepção de seus produtos e serviços no mercado.<sup>98</sup> Somente no ano de 2020, por exemplo, diversos fatores em âmbito internacional contribuíram para colocar a ética entre as prioridades das empresas de maneira inédita. A pandemia de Covid-19, as frequentes manifestações populares por justiça e equidade social, a aceleração da transformação digital em curso e a crescente tomada de consciência a respeito da discriminação e da exclusão de oportunidades para alguns segmentos da população intensificaram as demandas para que as organizações adotem certas medidas como resposta, que vão desde o apoio a causas sociais ao anúncio de incentivos para a promoção da diversidade. No mercado consumidor, por exemplo, uma pesquisa recente nos Estados Unidos revelou que a maioria

---

inevitable; stagnation serves no purpose, so how do we innovate responsibly? (...) The first step is to understand the role we play in all of this and to take responsibility for it. Think through the consequences as much as possible, both the direct and the unintended ones, then monitor for the reaction to it. Looking for your unknown unknowns is a really difficult position to be in but if you don't look you will probably never see them. At least monitoring what is happening gives you that fighting chance. Ultimately you have to take action. This is an evolving field, don't be afraid to admit it if things didn't go the way you wanted them to and you need to course-correct." CARRUTHERS, Caroline, JACKSON, Peter. Idem, *ibidem*. p. 139.

<sup>98</sup> Além das próprias vítimas diretas da discriminação algorítmica, qualquer descuido em relação ao uso ético da tecnologia pode rapidamente causar enormes estragos à imagem pública da empresa envolvida, prejudicando relações com clientes e entidades reguladoras, e reduzindo seus ganhos econômicos. Revelações recentes da Microsoft e do Google para grupos de investidores advertiram sobre o estrago em potencial que uma inteligência artificial utilizada de modo indevido pode causar em suas respectivas marcas. Mais detalhes estão disponíveis em: <https://www.theverge.com/2019/2/11/18220050/google-microsoft-ai-brand-damage-investors-10-k-filing>. Acesso em 30 de outubro de 2020.

significativa (68%) do público vê a sustentabilidade como sendo muito importante ao adquirir algum produto ou serviço, e percentual relativamente alto (49%) pagariam até mesmo a mais por produtos sustentáveis<sup>99</sup>. Esses dados são ainda mais significativos quando se considera as faixas etárias mais jovens. Os nascidos a partir da década de 1980 estão entre as pessoas que mais demonstram preocupação a respeito das consequências éticas da adoção de tecnologias envolvendo inteligência artificial em seus locais de trabalho.<sup>100</sup>

Em relação ao viés algorítmico especificamente, em pesquisa recente<sup>101</sup> entre profissionais de cargo gerencial em empresas da área de tecnologia, constatou-se que uma proporção bastante alta (94% nos Estados Unidos e 86% no Reino Unido) tinha a intenção de aumentar o investimento em medidas preventivas contra mecanismos de inteligência artificial enviesados no ano de 2020. Esse cenário se repete para outras categorias de trabalhadores da área de tecnologia. Um estudo<sup>102</sup> no Reino Unido identificou que 28% dos entrevistados presenciaram decisões de tecnologia eticamente questionáveis, conforme demonstra a figura 2 abaixo, e que 18% inclusive deixaram as organizações em que atuavam por conta deste fator.

O compromisso com padrões éticos mínimos demonstrado pelas empresas de tecnologia e seus profissionais se torna mais relevante em um cenário como o brasileiro, onde não há um sistema regulatório específico para lidar com o problema da discriminação algorítmica. Além disso, grande parte da população nacional não tem conhecimentos suficientes para utilizar os serviços tecnológicos de última geração, estando em posição vulnerável diante de quaisquer efeitos negativos que o uso das tecnologias mais recentes podem apresentar. A figura 3 a seguir demonstra como, em países menos desenvolvidos (como o Brasil), a dependência dos usuários em relação às principais empresas de tecnologia pode estar bastante acentuada, sobretudo no que diz respeito às redes sociais.<sup>103</sup> Este fato provavelmente contribui para que determinados vieses nessas plataformas sejam amplificados.<sup>104</sup>

---

<sup>99</sup> A pesquisa foi realizada pela empresa de tecnologia norte-americana CGS, e está disponível no seguinte endereço eletrônico: <https://www.cgsinc.com/en/infographics/CGS-Survey-Reveals-Sustainability-Is-Driving-Demand-and-Customer-Loyalty>. Acesso em 30 de outubro de 2020.

<sup>100</sup> A pesquisa foi realizada pela organização norte-americana Genesys em seis países de economia avançada. Detalhes disponíveis em: <https://www.prnewswire.com/news-releases/new-workplace-survey-finds-nearly-80-of-employers-arent-worried-about-unethical-use-of-ai-but-maybe-they-should-be-300911214.html>. Acesso em 30 de outubro de 2020.

<sup>101</sup> Maiores detalhes disponíveis em: <https://www.computerweekly.com/news/252474408/IT-chiefs-recognise-the-risks-of-artificial-intelligence-bias>. Acesso em 30 de outubro de 2020.

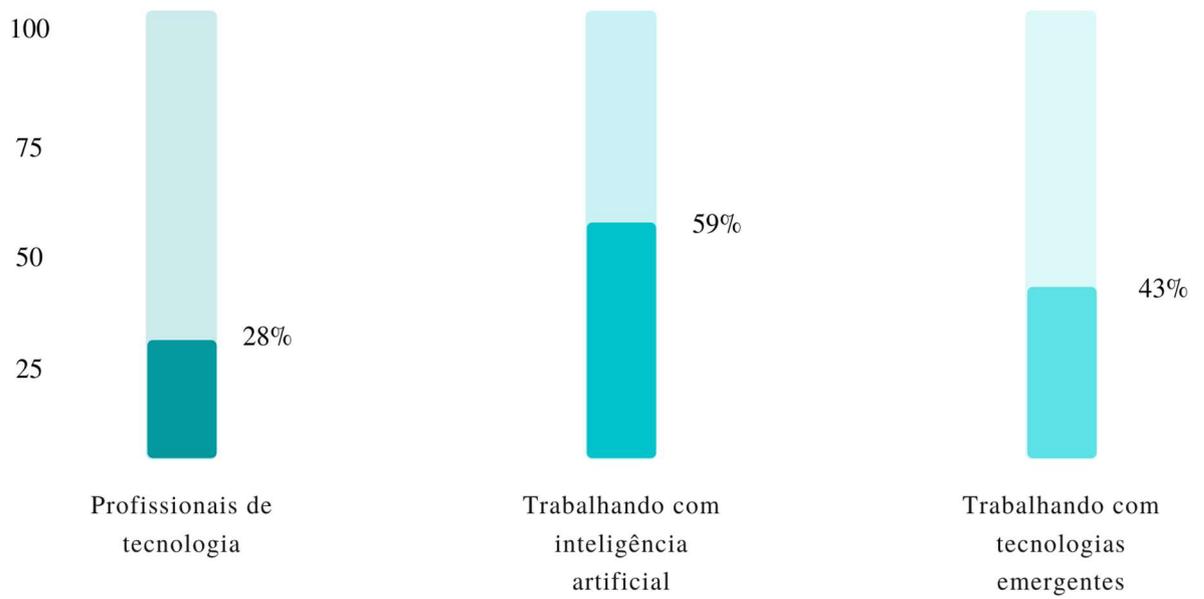
<sup>102</sup> O estudo da organização britânica Doteveryone pode ser consultado no seguinte endereço eletrônico: <https://www.doteveryone.org.uk/report/workersview/>. Acesso em 30 de outubro de 2020.

<sup>103</sup> Detalhes acerca da pesquisa estão disponíveis na página eletrônica a seguir: <https://qz.com/333313/millions-of-facebook-users-have-no-idea-theyre-using-the-internet/>. Acesso em 30 de outubro de 2020.

<sup>104</sup> Em 2018 o Facebook foi inclusive acusado pela ONU de negligência no caso do genocídio da minoria étnica Rohingya em Myanmar, no sudeste asiático. Detalhes em: <https://www.reuters.com/article/us-myanmar-rohingya-facebook-idUSKCN1GO2PN>. Acesso em 30 de outubro de 2020.

Figura 2

Proporção de profissionais de tecnologia que presenciaram decisões em seu local de trabalho que poderiam provocar consequências negativas para as pessoas e a sociedade

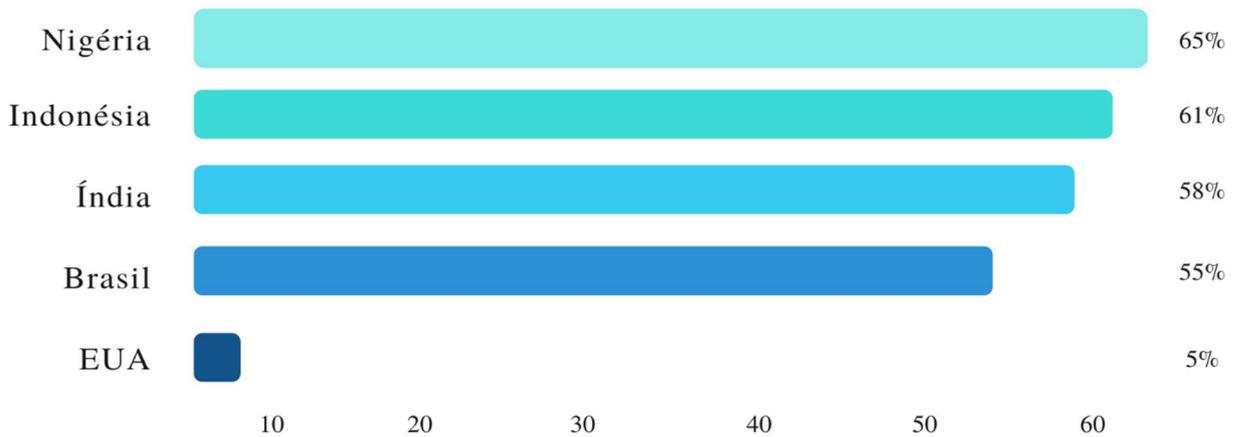


Fonte: Doteveryone

Figura 3

Percentual de entrevistados que concordam com a seguinte afirmação:

"O Facebook é a Internet"



Fonte: Quartz

Como este trabalho apresentou ao longo das páginas anteriores, existe uma série de desafios no uso ético de qualquer tecnologia, que deve ser administrada e monitorada tão ativamente quanto qualquer outro aspecto da instituição que faz uso da ferramenta em questão. No âmbito empresarial, as políticas internas de promoção de princípios éticos podem constituir o principal método para atingir esse objetivo e gerar resultados positivos para seus clientes e controladores.<sup>105</sup>

Na medida em que os mecanismos de inteligência artificial se desenvolvem e ganham maior complexidade, se torna mais difícil para qualquer organização avaliar se o seu funcionamento ocorre do modo pretendido. Este ponto é especialmente delicado em relação às áreas mais sensíveis, como assistência à saúde, justiça criminal e acesso a serviços financeiros. Em muitos casos pode mesmo ser difícil determinar qual medida seria apropriada diante de um dilema ético ou conflito de direitos fundamentais, o que torna ainda mais necessária a existência de uma resposta jurídica oficial para esse assunto. É neste ponto que a regulação jurídica propriamente dita pode desempenhar um papel complementar e bastante necessário para ir além das políticas internas organizacionais, como se demonstra a seguir. Até mesmo porque, a proteção aos direitos fundamentais não deve estar subordinada ao voluntarismo de quaisquer agentes do setor privado, devendo ser garantida pelo Estado por todos os meios disponíveis. Nenhuma reivindicação de conduta “ética” empresarial ou profissional pode ser considerada apta a substituir a urgência de restrições legalmente aplicáveis à implantação de tecnologias que detêm o potencial de violar as garantias básicas previstas no ordenamento jurídico brasileiro.

## 5.2 Regulação jurídica

O conceito de regulação pode apresentar diversos significados dependendo do contexto. Na Economia, por exemplo, uma das definições mais conhecidas é a do professor norte-americano Nicholas Gregory Mankiw, que afirma se tratar de uma intervenção estatal destinada a promover a eficiência e a equidade econômicas, isto é, resolver falhas de mercado em prol da competitividade e realizar justiça social.<sup>106</sup>

---

<sup>105</sup> Sobre este ponto, existem algumas propostas de princípios éticos gerais, como a formulada pela Comissão Europeia, disponível no seguinte endereço eletrônico: <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>. Há também outra proposta, de caráter similar, do Berkman Klein Center for Internet and Society da Universidade de Harvard, que está disponível aqui: <https://cyber.harvard.edu/topics/ethics-and-governance-ai>. Acesso em 30 de outubro de 2020.

<sup>106</sup> MANKIW, Nicholas Gregory. *Principles of Economics*. 5ª ed. South-Western Cengage Learning, 2007. p. 74.

Para o direito brasileiro há também uma variedade conceitual grande, onde algumas análises privilegiam um ou outro aspecto para caracterizar e classificar a regulação. Essa discussão doutrinária não será abordada aqui em virtude dos limites e propósitos deste trabalho. Entretanto, como se sabe, a noção jurídica não pode se limitar à simples reprodução de teorias políticas ou econômicas, ainda que de fato possua fundamentos políticos e econômicos. Sob o ponto de vista do direito administrativo especificamente, a regulação é exercida pelo Estado na prática por meio da edição de normas jurídicas em sentido amplo, ou seja, todas aquelas emitidas pelo Estado no exercício regular de suas funções legislativa, administrativa e jurisdicional. Neste sentido, de acordo com José dos Santos Carvalho Filho<sup>107</sup>, o poder público poderia atuar de duas maneiras na ordem econômica: enquanto “Estado Regulador” e “Estado Executor”. Nesta última hipótese haveria uma intervenção estatal específica para atuar diretamente na atividade econômica como se fizesse parte da iniciativa privada. Já a primeira categoria diz respeito ao poder de reger e fiscalizar a economia tal como é organizada pelos agentes econômicos privados, e é neste sentido que se analisa o papel do Estado ao longo desta pesquisa.

No caso brasileiro, a atividade reguladora do Estado está disciplinada entre os artigos 170 e 181 da Constituição de 1988. O artigo 170<sup>108</sup> estabelece as finalidades constitucionais da regulação, enquanto que o artigo 174<sup>109</sup> trata dos meios pelos quais ela é realizada: fiscalização, incentivo e planejamento<sup>110</sup>. A partir destes dois dispositivos é possível compreender que o Estado tem o poder-dever de agir positivamente para atingir os objetivos estabelecidos no texto constitucional no que tange à supervisão da ordem econômica, por meio do poder de polícia do qual dispõe. Deste modo, por exemplo, o poder público deve conciliar a manutenção da livre iniciativa com a proteção aos consumidores e a redução das desigualdades sociais, o que pode

---

<sup>107</sup> CARVALHO FILHO, José dos Santos. *Manual de Direito Administrativo*. 30ª ed. São Paulo: Atlas, 2016. p. 967.

<sup>108</sup> CRFB/88, Art. 170. A ordem econômica, fundada na valorização do trabalho humano e na livre iniciativa, tem por fim assegurar a todos existência digna, conforme os ditames da justiça social, observados os seguintes princípios: I - soberania nacional; II - propriedade privada; III - função social da propriedade; IV - livre concorrência; V - defesa do consumidor; VI - defesa do meio ambiente, inclusive mediante tratamento diferenciado conforme o impacto ambiental dos produtos e serviços e de seus processos de elaboração e prestação; VII - redução das desigualdades regionais e sociais; VIII - busca do pleno emprego; IX - tratamento favorecido para as empresas de pequeno porte constituídas sob as leis brasileiras e que tenham sua sede e administração no País. Parágrafo único. É assegurado a todos o livre exercício de qualquer atividade econômica, independentemente de autorização de órgãos públicos, salvo nos casos previstos em lei.

<sup>109</sup> CRFB/88, Art. 174. Como agente normativo e regulador da atividade econômica, o Estado exercerá, na forma da lei, as funções de fiscalização, incentivo e planejamento, sendo este determinante para o setor público e indicativo para o setor privado.

<sup>110</sup> A Lei nº 13.874/2019, conhecida como Lei da Liberdade Econômica não incluiu o direito do consumidor entre os ramos do direito abarcados por suas disposições, de modo que sua redação parece não trazer nenhuma determinação específica com consequências para os casos de discriminação algorítmica, em primeira análise.

ser feito por meio do regramento e inspeção específicos para cada atividade econômica. Neste sentido, o atual estágio tecnológico demanda ainda mais atenção e velocidade do legislador, considerando a complexidade dos “aspectos decorrentes da inovação na vida das pessoas e das empresas, desde os direitos de personalidade e intimidade, até aqueles pertinentes ao âmbito do direito do consumidor, tributário, dos contratos, criminal, propriedade imaterial e outros.”<sup>111</sup>

Se no Brasil ainda não há política regulatória específica, no direito estrangeiro a União Europeia tem sido o principal fórum de discussão política e de regulação do *big data*.<sup>112</sup> Conforme indicado na parte inicial deste texto, o RGPD prevê especificamente a revisão humana de decisões automatizadas como um direito dos indivíduos. Mas ainda permanecem desafios, e o Parlamento Europeu deve seguir implementando novas normas que regulem a área de tecnologia em seus diferentes aspectos.<sup>113</sup>

Na visão de Marietje Schaake,<sup>114</sup> que atuou como deputada no Parlamento Europeu na última década, a União Europeia seria a única reguladora de fato das grandes empresas transnacionais de tecnologia sediadas principalmente no Vale do Silício (no estado norte-americano da Califórnia). Seu principal argumento é de que a tecnologia só pode ser democrática quando submetida a um debate público envolvendo os cidadãos e seus representantes políticos, de modo que a tributação e a regulação dessa área econômica não significam um ataque às empresas de informática nem aos seus produtos. O objetivo da regulação consiste em garantir que os produtos sejam seguros e estejam em conformidade com

---

<sup>111</sup> DOMINGUES, Paulo Sérgio. *Legislativo 4.0: o desafio da criação de novas leis para um mundo em mutação*. In: Cadernos Adenauer XXI, nº1. A quarta revolução industrial: inovações, desafios e oportunidades. Rio de Janeiro: Fundação Konrad Adenauer, 2020. p. 56. Disponível em: <https://www.kas.de/pt/web/brasilien/einzeltitel/-/content/cadernos-adenauer-1-2020-1>. Acesso em 01 de novembro de 2020.

<sup>112</sup> Nos Estados Unidos a *California Consumer Privacy Act* é a principal lei que regula o uso dos dados, enquanto não é editada norma federal que trate do assunto. Sobre este assunto: HARTZOG, Woodrow; RICHARDS, Neil. *Privacy's Constitutional Moment and the Limits of Data Protection*, 61 B.C.L. Rev. 1687 (2020). Disponível em: <https://lawdigitalcommons.bc.edu/bclr/vol61/iss5/3>. Acesso em 29 de novembro de 2020.

<sup>113</sup> No cenário internacional as organizações sem fins lucrativos Anistia Internacional e Access Now formularam a “Declaração de Toronto”, um documento que tem por objetivo estabelecer parâmetros para a operação de algoritmos de aprendizado de máquina, e convidando Estados e empresas de tecnologia a garantir que os algoritmos respeitem os princípios básicos de igualdade e não-discriminação. O documento pode ser acessado aqui: [https://www.accessnow.org/cms/assets/uploads/2018/08/The-Toronto-Declaration\\_ENG\\_08-2018.pdf](https://www.accessnow.org/cms/assets/uploads/2018/08/The-Toronto-Declaration_ENG_08-2018.pdf). Acesso em 31 de outubro de 2020.

<sup>114</sup> Conforme entrevistas disponibilizadas nos seguintes endereços eletrônicos: <https://www.foreignaffairs.com/articles/world/2020-10-13/lawless-realm>; [https://www.newyorker.com/tech/annals-of-technology/what-can-america-learn-from-europe-about-regulating-big-tech?utm\\_source=twitter&utm\\_medium=social&utm\\_campaign=onsite-share&utm\\_brand=the-new-yorker&utm\\_social-type=earned](https://www.newyorker.com/tech/annals-of-technology/what-can-america-learn-from-europe-about-regulating-big-tech?utm_source=twitter&utm_medium=social&utm_campaign=onsite-share&utm_brand=the-new-yorker&utm_social-type=earned). Acesso em 31 de outubro de 2020.

as leis e os valores das sociedades em que operam, observando em especial os critérios de transparência e privacidade.<sup>115</sup>

Segundo Schaake, os segredos comerciais e os acordos de confidencialidade muitas vezes impedem que as informações sobre o funcionamento das grandes empresas privadas de tecnologia sejam divulgadas ao público. Essas proteções jurídicas dificultam a resposta estatal em relação aos riscos que já existem e também impedem pesquisas independentes sobre os efeitos pretendidos e não intencionais dos produtos dessas organizações. O resultado é o bloqueio do debate público bem informado sobre digitalização e segurança, inibindo a formulação de políticas com base em evidências. Para Schaake, o Estado tem o dever de regular esta seara para garantir o acesso às informações de interesse público, lembrando que a iniciativa privada deve atuar em consonância com os princípios universais dos direitos humanos, o que precisa ser exigido das lideranças políticas internacionais.

Conforme indicado, no âmbito europeu o RGPD disciplinou a tomada de decisão automatizada na era do *big data* e estabeleceu a revisão humana como um meio de garantir a transparência de tais decisões. Além disso, no caso de decisões automatizadas envolvendo dados pessoais, o RGPD obriga o responsável pelo tratamento a fornecer ao titular dos dados "informações significativas sobre a lógica envolvida", conforme se depreende da leitura do artigo 13, inciso 2, alínea f e artigo 14, inciso 2, alínea g.<sup>116</sup> Neste ponto, existe um debate a respeito do direito do usuário de obter explicações sobre a decisão automatizada que afete seus interesses jurídicos. Embora esse direito em princípio se encaixe bem na intenção mais ampla do RGPD de promover um alto nível de transparência, ele também levanta algumas questões, como o que exatamente precisaria ser revelado ao titular dos dados e o que se entende por explicação de uma decisão automatizada baseada em algoritmos. Além de obstáculos técnicos, também existem questões relacionadas à propriedade intelectual que não podem ser

---

<sup>115</sup> No que tange a privacidade, o professor norte-americano Daniel J. Solove argumenta que é mais eficaz criar mecanismos regulatórios com foco no acesso, uso, armazenamento e transferência de dados do que instituir direitos (e responsabilidades) aos usuários (detentores destes dados). SOLOVE, Daniel J., *The Myth of the Privacy Paradox*. George Washington Law Review, Vol. 89, 2021. p. 40. Disponível em: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3536265](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3536265). Acesso em 29 de novembro de 2020.

<sup>116</sup> RGPD, artigo 13: informações a facultar quando os dados pessoais são recolhidos junto do titular: (...) 2. Para além das informações referidas no nº 1, aquando da recolha dos dados pessoais, o responsável pelo tratamento fornece ao titular as seguintes informações adicionais, necessárias para garantir um tratamento equitativo e transparente: (...) f) A existência de decisões automatizadas, incluindo a definição de perfis, referida no artigo 22.o, nºs 1 e 4, e, pelo menos nesses casos, informações úteis relativas à lógica subjacente, bem como a importância e as consequências previstas de tal tratamento para o titular dos dados.; O artigo 14 possui idêntica determinação, porém incidindo em relação aos casos em que as informações não são recolhidas junto ao seu titular.

negligenciadas.<sup>117</sup> Nos próximos anos, o Parlamento Europeu deverá enfrentar o desafio de buscar encontrar formas de conciliar a exigência de transparência com os interesses das empresas de tecnologia.

O debate se torna ainda mais complexo na medida em que há posições contrárias ao direito de revisão humana das decisões automatizadas (abordada no capítulo 3). Existem pesquisadores<sup>118</sup> que defendem que a intervenção humana em um processo automatizado e o direito à informação a respeito de como a decisão foi tomada pelo algoritmo dificilmente serão remédios jurídicos suficientes para os danos causados pela inteligência artificial, como a discriminação algorítmica. Nesse raciocínio, por exemplo, no caso do RGPD ainda não haveria clareza a respeito dos institutos elencados em seu texto; tampouco existiria uma definição precisa sobre o tipo de informação a respeito do funcionamento da tecnologia que deve ser disponibilizada ao usuário. Neste caso, a solução estaria em outros dispositivos trazidos pelo próprio RGPD, como o direito ao esquecimento e a portabilidade de dados pessoais, bem como processos de avaliação e certificação dos sistemas algorítmicos por instituições credenciadas.

Assim, a regulação jurídica da inteligência artificial deve ter o objetivo de proteger os usuários de eventuais danos e manter o mercado livre para a competição econômica, ou seja, preservar o interesse público e proteger o desenvolvimento tecnológico. Qualquer regulamento nesse sentido deve sopesar os riscos e os benefícios que a adoção da inteligência artificial pode trazer para a sociedade. Inclusive, as leis já existentes deverão ser adaptadas a essa nova realidade, em que a tecnologia se torna cada vez mais relevante para o atual modelo econômico global. Além disso, os reguladores devem ter em mente que seu trabalho não pode restringir a inovação tecnológica benéfica em virtude do aumento exagerado dos custos de conformidade com as normas e parâmetros estabelecidos. O governo federal norte-americano apontou que a regulação da inteligência artificial poderia inclusive ser pensada caso a caso, sendo inserida nas normas que já regem determinado produto no momento em que o mesmo seja atualizado para

---

<sup>117</sup> Na dinâmica do atual desenvolvimento tecnológico é válido refletir sobre certa limitação nos direitos de propriedade industrial, tomando por base a teoria da função social da propriedade. Essa medida pode inclusive auxiliar o desenvolvimento dos países menos desenvolvidos como o Brasil diante dos atuais desafios trazidos pela globalização. Sobre este assunto, consultar: BRUCH, Kelly Lissandra; HOFF, Debora Nayar; DEWES, Homero. *A função social do direito de propriedade industrial como alternativa de governabilidade aos países em desenvolvimento: um estudo sobre a propriedade industrial de plantas*. Direito, Estado e Sociedade, v. 32, 2008. p. 149. Disponível em:

[https://www.researchgate.net/publication/236897426\\_Funcao\\_social\\_do\\_direito\\_de\\_propriedade\\_industrial\\_com\\_o\\_alternativa\\_de\\_governabilidade\\_ao\\_paises\\_em\\_desenvolvimento\\_um\\_estudo\\_sobre\\_a\\_propriedade\\_industrial\\_de\\_plantas](https://www.researchgate.net/publication/236897426_Funcao_social_do_direito_de_propriedade_industrial_com_o_alternativa_de_governabilidade_ao_paises_em_desenvolvimento_um_estudo_sobre_a_propriedade_industrial_de_plantas). Acesso em 30 de outubro de 2020.

<sup>118</sup> Sobre este assunto, consultar: EDWARDS, Lilian; VEALE, Michael. *Slave to the Algorithm? Why a Right to Explanation is Probably Not the Remedy You are Looking for*. SSRN Electronic Journal, 2017. Disponível em: [https://www.researchgate.net/publication/318003528\\_Slave\\_to\\_the\\_Algorithm\\_Why\\_a\\_Right\\_to\\_Explanation\\_is\\_Probably\\_Not\\_the\\_Remedy\\_You\\_are\\_Looking\\_for](https://www.researchgate.net/publication/318003528_Slave_to_the_Algorithm_Why_a_Right_to_Explanation_is_Probably_Not_the_Remedy_You_are_Looking_for). Acesso em 01 de novembro de 2020.

incluir alguma ferramenta inteligente (como por exemplo no caso da indústria automobilística em relação aos carros autônomos).<sup>119</sup>

Como se sabe, a regulamentação eficaz de qualquer tecnologia complexa pressupõe uma assessoria técnica especializada para ajudar a orientar a tomada de decisão dos representantes políticos em seu trabalho legislativo, além de supervisionar todas as fases do processo regulatório. Esse é um desafio ainda maior para países em desenvolvimento como o Brasil, onde provavelmente será necessário desenvolver pesquisas e produzir conhecimento a respeito das particularidades da realidade nacional diante das potencialidades e riscos que a inteligência artificial pode representar. O trabalho científico realizado por profissionais de diferentes perspectivas e qualificações, a participação pública e a troca de conhecimento entre o governo federal e os governos estaduais e municipais, universidades, centros de pesquisa, dentre outras organizações pode contribuir para que o melhor modelo regulatório seja alcançado.

Talvez o principal objetivo da regulação seja fomentar a criação de melhores mecanismos de inteligência artificial desde o princípio, para que não seja preciso se dedicar tanto à salvaguarda de direitos individuais de cada usuário. Além disso, o fortalecimento de agências reguladoras e organizações da sociedade civil pode ajudar na supervisão e análise dos critérios de precisão, imparcialidade e integridade que se espera das novas tecnologias.

Em síntese, este capítulo abordou a ética e a regulação como maneiras de mitigar os efeitos da discriminação algorítmica, tarefa que deve ser interdisciplinar e envolver as empresas de tecnologia (e outras organizações da sociedade civil) e o poder público. Como indicado ao longo do texto, este é um desafio significativo. Porém, o esforço de evitar a discriminação algorítmica e proteger os indivíduos de seus resultados prejudiciais indesejados é uma medida que pode gerar a confiança de que a utilização da inteligência artificial é segura e fomentar ainda mais o desenvolvimento dessa modalidade tecnológica.

---

<sup>119</sup> THE WHITE HOUSE. *Preparing for the Future of Artificial Intelligence*. Executive Office of the President, 2016. pp. 17. Disponível em: [https://obamawhitehouse.archives.gov/sites/default/files/whitehouse\\_files/microsites/ostp/NSTC/preparing\\_for\\_the\\_future\\_of\\_ai.pdf](https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf). Acesso em 01 de novembro de 2020.

## 6 Conclusão

A motivação para esta pesquisa se deve ao fato de que a discussão sobre os impactos sociais e jurídicos das novas ferramentas tecnológicas disponíveis atualmente são fundamentais para que o Direito consiga contribuir com as soluções adequadas que a sociedade espera. É preciso evitar que os operadores do Direito não apenas venham a cancelar narrativas de determinadas áreas ou sujeitos e possam, ao contrário, permitir maior diversidade de perspectivas que atendam aos anseios plurais de uma democracia. Como a aplicação da legislação e da jurisprudência depende do intérprete e a hermenêutica exerce grande influência no Direito, talvez seja válida a analogia em relação à revisão humana de decisões automatizadas, enquanto proposta apta para mitigar eventuais vieses em sistemas de inteligência artificial. Entretanto, ainda não há consenso científico a respeito da capacidade dessa medida de resolver os problemas gerados pela discriminação algorítmica.

Conforme demonstrado ao longo deste trabalho, os algoritmos em sistemas de *big data* são treinados e alimentados por um conjunto gigantesco de dados, estando muito além do que o ser humano conseguiria lidar de modo solitário. Porém, se esses dados são tendenciosos ou intencionalmente distorcidos, os algoritmos se tornarão conseqüentemente parciais em sua operação. Apesar de processarem uma grande quantidade de informações com muita velocidade, esses sistemas eletrônicos não são infalíveis nem se encontram livres da influência humana.

Não se trata aqui de negar que a inteligência artificial já trouxe grandes benefícios para a sociedade. Existe a crença inclusive de que seu uso torne o mundo mais justo e igualitário, conforme demonstrado ao longo do texto. A regulação jurídica tampouco deve travar o avanço científico ou dificultar a inovação. Seu propósito é apenas garantir que os direitos fundamentais não sejam colocados em segundo plano em favor de determinada agenda econômica. A dificuldade reside na velocidade das mudanças tecnológicas em curso, o que requer trabalho e atenção constante por parte do poder público. Promover critérios de transparência pode ajudar no controle democrático da tecnologia, conforme visto. Mas somente será possível obter a abordagem multissetorial, multidisciplinar e conjunta que esta questão requer mediante a dedicação de especialistas em diversas áreas, como direito, ciência da computação e ética, além de instituições como órgãos públicos, academia, sociedade civil e empresas privadas.

Os conceitos abordados no texto mostram interseção cada vez maior entre as áreas do direito e da ciência da computação. Neste sentido, a pesquisa realizou breve exposição a respeito de alguns dos mecanismos mais comuns de inteligência artificial, além de eventuais

tipos de vieses que eles podem apresentar. Na falta de solução mais adequada para este problema, a pergunta deste trabalho abrange a revisão humana das operações algorítmicas como um meio de correção para os eventuais resultados enviesados que forem entregues por sistemas automatizados. Ao deixar de prever a revisão humana expressamente, a legislação brasileira é percebida aqui como insuficiente para lidar com esta situação. Acredita-se que a regulação deveria ser capaz de fomentar o desenvolvimento tecnológico com atenção aos direitos fundamentais e conquistas sociais. Este pode ser o cenário mais favorável para que a tecnologia de fato consiga contribuir com a melhoria das condições de vida ao longo deste século.

É inevitável que uma pesquisa realizada sobre algoritmos de inteligência artificial já esteja desatualizada no momento de sua publicação, em decorrência do desenvolvimento bastante rápido que essas ferramentas apresentam atualmente. Entretanto, é de se esperar que os desafios abordados ao longo deste trabalho continuem existindo por mais algum tempo. Neste sentido, o principal propósito do texto permanece como um convite ao debate sobre alguns dos efeitos indesejáveis que o desenvolvimento tecnológico pode trazer aos indivíduos, para que em seguida seja possível encontrar soluções eficazes. No Brasil, em especial, a História demonstra que as profundas desigualdades sociais não desaparecerão sem que exista engajamento e resistência por parte da sociedade.

Este trabalho se encerra como uma reflexão a respeito das alterações relevantes que os algoritmos e suas aplicações tem causado nas relações sociais, políticas e econômicas contemporâneas, para as quais se acredita que a sociedade brasileira estaria melhor amparada se houvesse uma política pública especialmente elaborada para lidar com este tema. Essa poderia ser uma resposta efetiva para conciliar eventual conflito de direitos fundamentais envolvidos nessa questão, como liberdade econômica e igualdade. O desafio principal da regulação jurídica aqui é encontrar a melhor maneira de proteger a inovação e os direitos humanos, não só no Brasil como no restante do mundo.

## REFERÊNCIAS

ABE, Jair Minoro; SCALZITTI, Alexandre; SILVA FILHO, João Inácio. *Introdução à lógica para a ciência da computação*. São Paulo: Arte e Ciência, 2002.

ARANHA, Maria Lúcia de Arruda; MARTINS, Maria Helena Pires. *Filosofando* : Introdução à Filosofia. São Paulo : Moderna, 1986.

BATES, Madeleine. *Models of natural language understanding*. Proceedings of the National Academy of Sciences, v. 92, n. 22, 1995. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC40721/pdf/pnas01500-0075.pdf> Acesso em 23 de agosto de 2020.

BARROSO, Luís Roberto. *Diferentes, mas iguais: o reconhecimento jurídico das relações homoafetivas no Brasil*. Revista Brasileira de Direito Constitucional, v. 17, p. 105-138, 2011. Disponível em: [http://www.luisrobertobarroso.com.br/wp-content/uploads/2017/09/diferentes\\_mas\\_iguais\\_atualizacao\\_2011.pdf](http://www.luisrobertobarroso.com.br/wp-content/uploads/2017/09/diferentes_mas_iguais_atualizacao_2011.pdf). Acesso em 26 de outubro de 2020

BITTAR, Eduardo Carlos Bianca. *Curso de ética jurídica: ética geral e profissional*. 13ª ed. São Paulo: Saraiva, 2016.

BRASIL. Constituição da República Federativa do Brasil. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/constituicao/constituicao.htm](http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm). Acesso em: 12 de novembro de 2020.

\_\_\_\_\_. Lei nº 9.609 , de 19 de fevereiro de 1998. Dispõe sobre a proteção da propriedade intelectual de programa de computador, sua comercialização no País, e dá outras providências. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/leis/19609.htm](http://www.planalto.gov.br/ccivil_03/leis/19609.htm). Acesso em 30 de outubro de 2020.

\_\_\_\_\_. Lei n.º 9.610, de 16 de fevereiro de 1996. Altera, atualiza e consolida a legislação sobre direitos autorais e dá outras providências. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/leis/19610.htm](http://www.planalto.gov.br/ccivil_03/leis/19610.htm). Acesso em 30 de outubro de 2020.

\_\_\_\_\_. Lei nº 10.406, de 10 de janeiro de 2002. Institui o Código Civil. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/leis/2002/110406compilada.htm](http://www.planalto.gov.br/ccivil_03/leis/2002/110406compilada.htm). Acesso em 12 de novembro de 2020.

\_\_\_\_\_. Lei nº 12.965, de 23 de abril de 2014. Estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/\\_ato2011-2014/2014/lei/112965.htm](http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/112965.htm). Acesso em 17 de agosto de 2020.

\_\_\_\_\_. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). Disponível em: [http://www.planalto.gov.br/ccivil\\_03/\\_ato2015-2018/2018/lei/L13709.htm](http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm). Acesso em 12 de novembro de 2020.

\_\_\_\_\_. Lei nº 13.874, de 20 de setembro de 2019. Institui a Declaração de Direitos de Liberdade Econômica; estabelece garantias de livre mercado; altera as Leis nos 10.406, de 10

de janeiro de 2002 (Código Civil), 6.404, de 15 de dezembro de 1976, 11.598, de 3 de dezembro de 2007, 12.682, de 9 de julho de 2012, 6.015, de 31 de dezembro de 1973, 10.522, de 19 de julho de 2002, 8.934, de 18 de novembro 1994, o Decreto-Lei nº 9.760, de 5 de setembro de 1946 e a Consolidação das Leis do Trabalho, aprovada pelo Decreto-Lei nº 5.452, de 1º de maio de 1943; revoga a Lei Delegada nº 4, de 26 de setembro de 1962, a Lei nº 11.887, de 24 de dezembro de 2008, e dispositivos do Decreto-Lei nº 73, de 21 de novembro de 1966; e dá outras providências. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/\\_ato2019-2022/2019/lei/L13874.htm](http://www.planalto.gov.br/ccivil_03/_ato2019-2022/2019/lei/L13874.htm). Acesso em 30 de outubro de 2020.

\_\_\_\_\_. Provimento nº 71 da Corregedoria Nacional de Justiça, de 13 de junho de 2018. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/2608>. Acesso em 30 de outubro de 2020.

\_\_\_\_\_. Projeto de Lei nº 4496, de 2019. Altera a Lei nº 13.709, de 14 de agosto de 2018 (Lei Geral de Proteção de Dados Pessoais – LGPD), para definir a expressão “decisão automatizada”. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/138136>. Acesso em 12 de agosto de 2020.

\_\_\_\_\_. Resolução nº 332/2020 do Conselho Nacional de Justiça. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3429>. Acesso em 29 de novembro de 2020.

BROUSSARD, Meredith. *Artificial unintelligence: how computers misunderstand the world*. Cambridge: MIT Press, 2018.

BRUCE, Andrew; BRUCE, Peter. *Practical Statistics for Data Scientists*. Sebastopol: O’Reilly, 2017.

BRUCH, Kelly Lissandra; HOFF, Debora Nayar; DEWES, Homero. *A função social do direito de propriedade industrial como alternativa de governabilidade aos países em desenvolvimento: um estudo sobre a propriedade industrial de plantas*. Direito, Estado e Sociedade, v. 32, 2008. Disponível em: [https://www.researchgate.net/publication/236897426\\_Funcao\\_social\\_do\\_direito\\_de\\_propriedade\\_industrial\\_como\\_alternativa\\_de\\_governabilidade\\_ao\\_s\\_paises\\_em\\_desenvolvimento\\_um\\_estudo\\_sobre\\_a\\_propriedade\\_industrial\\_de\\_plantas](https://www.researchgate.net/publication/236897426_Funcao_social_do_direito_de_propriedade_industrial_como_alternativa_de_governabilidade_ao_s_paises_em_desenvolvimento_um_estudo_sobre_a_propriedade_industrial_de_plantas). Acesso em 30 de outubro de 2020.

CARRUTHERS, Caroline, JACKSON, Peter.; *The Chief Data Officer’s Playbook*. EBSCO Publishing : eBook Collection, 2018.

CARVALHO FILHO, José dos Santos. *Manual de Direito Administrativo*. 30ª ed. São Paulo: Atlas, 2016.

CARVALHO, Patrícia Heloisa de. *O Marco Civil da Internet: Uma análise sobre a constitucionalidade do artigo 19*. Revista da Faculdade de Direito do Sul de Minas. Pouso Alegre, 2017, v. 33, p. 228-244. Disponível em: <https://www.fdsu.edu.br/adm/artigos/6917c36392274c9b6393c7f7a7bddd1.pdf>. Acesso em 18 de agosto de 2020.

COPELAND, B. Jack, *The Modern History of Computing*. The Stanford Encyclopedia of Philosophy (Winter 2017 Edition), Edward N. Zalta (ed.). Disponível em: <https://plato.stanford.edu/entries/computing-history/>. Acesso em 3 de julho de 2020

DATTA, A., TSCHANTZ, M. C., *Automated experiments on ad privacy settings: A tale of opacity, choice, and discrimination*. In Proceedings on Privacy Enhancing Technologies (PoPETs). 2015. <https://www.andrew.cmu.edu/user/danupam/dtd-pets15.pdf>. Acesso em 19 de outubro de 2020.

DIMOULIS, Dimitri; MARTINS, Leonardo. *Teoria geral dos direitos fundamentais* – 5. ed. rev., atual. e ampl. – São Paulo: Atlas, 2014.

DOMINGUES, Paulo Sérgio. *Legislativo 4.0: o desafio da criação de novas leis para um mundo em mutação*. In: Cadernos Adenauer XXI, nº1. A quarta revolução industrial: inovações, desafios e oportunidades. Rio de Janeiro: Fundação Konrad Adenauer, 2020. Disponível em: <https://www.kas.de/pt/web/brasilien/einzelitel/-/content/cadernos-adenauer-1-2020-1>. Acesso em 01 de novembro de 2020.

DUTRA, Luciano. *Direito Constitucional Essencial*. Rio de Janeiro: Forense. 2ªed. 2016.

EDWARDS, Lilian; VEALE, Michael. *Slave to the Algorithm? Why a Right to Explanation is Probably Not the Remedy You are Looking for*. SSRN Electronic Journal, 2017. Disponível em: [https://www.researchgate.net/publication/318003528\\_Slave\\_to\\_the\\_Algorithm\\_Why\\_a\\_Right\\_to\\_Explationn\\_is\\_Probably\\_Not\\_the\\_Remedy\\_You\\_are\\_Looking\\_for](https://www.researchgate.net/publication/318003528_Slave_to_the_Algorithm_Why_a_Right_to_Explationn_is_Probably_Not_the_Remedy_You_are_Looking_for). Acesso em 01 de novembro de 2020.

ELIAS, Norbert. *Os estabelecidos e os outsiders: sociologia das relações de poder a partir de uma pequena comunidade*. Rio de Janeiro: Jorge Zahar, 2000.

FRANÇA, Phillip Gil. *Objetivos Fundamentais da República, Escolhas Públicas e Políticas Públicas: Caminhos de Concretização dos Benefícios Sociais Constitucionais*. Direitos sociais e políticas públicas I. 1ed. Curitiba: Clássica Editora, 2014, v. 25.

HARTZOG, Woodrow; RICHARDS, Neil. *Privacy's Constitutional Moment and the Limits of Data Protection*, 61 B.C.L. Rev. 1687 (2020). Disponível em: <https://lawdigitalcommons.bc.edu/bclr/vol61/iss5/3>. Acesso em 29 de novembro de 2020.

HASTIE, Trevor; TIBSHIRANI, Robert; FRIEDMAN, J. H. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2ª ed. Nova York: Springer, 2009. Disponível em: <https://link.springer.com/book/10.1007%2F978-0-387-84858-7#authorsandaffiliationsbook>. Acesso em 6 de junho de 2020.

HURWITZ, Judith; KIRSCH, Daniel. *Machine Learning For Dummies*: IBM Limited Edition. Hoboken: John Wiley & Sons Inc., 2018. Disponível em: <https://www.ibm.com/downloads/cas/GB8ZMQZ3>. Acesso em 16 de março de 2020.

LAWRENCE, B. Solum. *Legal Personhood for Artificial Intelligences*. North Carolina Law Review, v. 70, n. 4, 1992. p. 1236. Disponível em: <https://scholarship.law.unc.edu/cgi/viewcontent.cgi?article=3447&context=nclr>. Acesso em 16 de março de 2020.

MACHADO, Anderson Fraiha; MENDES, Leilani Dian; GARCEZ, Lucas Nogueira. *Introdução à Arquitetura e Engenharia Jurídica com Lawtex*. São Paulo: Looplex, 2018.

MACHADO, Eduarda Sordi Pinheiro. *Inteligência artificial e direitos autorais: a proteção de obras criadas por computadores inteligentes*. Trabalho de conclusão de curso de graduação. Universidade Federal do Rio Grande do Sul, Porto Alegre, 2019.

MANKIW, Nicholas Gregory. *Principles of Economics*. 5ª ed. South-Western Cengage Learning, 2007.

MATTHEWS, P. H. *Linguistics: a very short introduction*. Oxford: Oxford University Press, 2003.

MCCORDUCK, Pamela. *Machines who think : a personal inquiry into the history and prospects of artificial intelligence*. AK Peters Ltd : Natick, 2004

MOLNAR, Adam. *Technology, law, and the formation of (il)liberal democracy?*, Surveillance and society, vol. 15, no. 3/4, 2017, pp. 318-388. Disponível em <https://ojs.library.queensu.ca/index.php/surveillance-and-society/article/view/6645>. Acesso em 3 de junho de 2020.

NOBLE, Safiya Umoja. *Algorithms of oppression: how search engines reinforce racism*. New York: New York University Press, 2018.

OBERMEYER, Ziad; POWERS, Brian; VOGELI, Christine; MULLAINATHAN, Sendhil. *Dissecting racial bias in an algorithm used to manage the health of populations*. Science n° 366, 2019. pp. 447-453. Disponível em: <https://www.ehdc.org/sites/default/files/resources/files/Dissecting%20racial%20bias%20in%20an%20algorithm%20used%20to%20manage%20the%20health%20of%20populations.pdf>. Acesso em 22 de outubro de 2020.

O'NEIL, Cathy. *Weapons of math destruction: how big data increases inequality and threatens democracy*. New York: Crown Publishers, 2016.

PEARL, Judea. *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press, 2000. Disponível em: <http://bayes.cs.ucla.edu/BOOK-99/book-toc.html>. Acesso em 19 de outubro de 2020.

PRESTES, Edson. *Teoria dos grafos*. Porto Alegre: 2011, vol. 4, 2016. Disponível em: <http://www.inf.ufgrs.br/~prestes/Courses/Graph%20Theory/Livro/ParteLivroGrafos.pdf>. Acesso em 16 de março de 2020.

SOLOVE, Daniel J., The Myth of the Privacy Paradox. George Washington Law Review, Vol. 89, 2021. Disponível em: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3536265](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3536265). Acesso em 29 de novembro de 2020.

SHANNON, Claude Elwood. *A symbolic analysis of relay and switching circuits*. In Transactions of the American Institute of Electrical Engineers, vol. 57, no. 12, pp. 713-723, Dec. 1938. Disponível em: <https://www.cs.virginia.edu/~evans/greatworks/shannon38.pdf>. Acesso em 4 de julho de 2020.

SILVA, Tarcízio. *Racismo Algorítmico em Plataformas Digitais: microagressões e discriminação em código*. In: Anais do IV Simpósio Internacional LAVITS – Assimetrias e (In)visibilidades: Vigilância, Gênero e Raça. Salvador, Bahia, Brasil, 2019. Disponível em: [https://www.researchgate.net/publication/333700308\\_Racismo\\_Algoritmico\\_em\\_Plataformas\\_Digitais\\_microagressoes\\_e\\_discriminacao\\_em\\_codigo](https://www.researchgate.net/publication/333700308_Racismo_Algoritmico_em_Plataformas_Digitais_microagressoes_e_discriminacao_em_codigo). Acesso em 25 de outubro de 2020.

SOUZA, Carlos Affonso Pereira de; SILVA JUNIOR, Ronaldo Lemos da. *Marco Civil da Internet: construção e aplicação*. 1. ed. Juiz de Fora: Editar, 2016. v. 1. p. 25. Disponível em: < [https://itsrio.org/wp-content/uploads/2017/02/marco\\_civil\\_construcao\\_aplicacao.pdf](https://itsrio.org/wp-content/uploads/2017/02/marco_civil_construcao_aplicacao.pdf)>. Acesso em 18 de agosto de 2020

STANKOVIĆ; Radomir S.; ASTOLA, Jaakko. *Reprints from the Early Days of Information Sciences: TICSP Series On the Contributions of Akira Nakashima to Switching Theory*. Tampere: Tampere International Center for Signal Processing, 2008. Disponível em: <http://ticsp.cs.tut.fi/reports/reprint-nakashima-rr.pdf>. Acesso em 4 de julho de 2020.

STEFFENS, Luana. *A influência dos vieses cognitivos na tomada de decisão pela inteligência artificial: um estudo baseado nas evidências no caso norte-americano COMPAS*. In: Fabio da Silva Veiga; Denise Pires Fincato. (Org.). *Estudos de Direito, Desenvolvimento e novas Tecnologias*. 1ed. Porto Alegre: Instituto Iberoamericano de Estudos Jurídicos, 2020, v. 1, p. 1-419.

SURDEN, Harry. *Machine Learning and Law*. *Washington Law Review*, v. 89, 2014. p. 89. Disponível em: <https://digitalcommons.law.uw.edu/wlr/vol89/iss1/5/> Acesso em 16 de março de 2020.

SWEENEY, L. *Discrimination in Online Ad Delivery*. Harvard University. January 28, 2013 <http://ssrn.com/abstract=2208240>. Acesso em 19 de outubro de 2020.

THE WHITE HOUSE. *Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights*. Executive Office of the President, 2016. Disponível em: [https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/2016\\_0504\\_data\\_discrimination.pdf](https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf). Acesso em 30 de agosto de 2020.

\_\_\_\_\_. *Preparing for the Future of Artificial Intelligence*. Executive Office of the President, 2016. pp. 17. Disponível em: [https://obamawhitehouse.archives.gov/sites/default/files/whitehouse\\_files/microsites/ostp/NS-TC/preparing\\_for\\_the\\_future\\_of\\_ai.pdf](https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NS-TC/preparing_for_the_future_of_ai.pdf). Acesso em 01 de novembro de 2020.

TURING, Allan M. *Computing Machinery and Intelligence*. *Mind*, v. 59, n. 236, 1950. Disponível em: <https://academic.oup.com/mind/article-abstract/LIX/236/433/986238>. Acesso em 29 de fevereiro de 2020.

UNIÃO EUROPEIA. Regulamento 2016/679 do Parlamento Europeu e do Conselho. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:32016R0679&from=PT#d1e40-1-1>. Acesso em 16 de março de 2020.

WANG, Yilun.; KOSINSKI, Michal. *Deep neural networks are more accurate than humans at detecting sexual orientation from facial images*. *Journal of Personality and Social Psychology*.

American Psychological Association, 2017. Disponível em: <<https://osf.io/zn79k/>>. Acesso em 6 de junho de 2020

WILSON, Benjamin; HOFFMAN, Judy; MORGENSTERN, Jamie. *Predictive Inequity in Object Detection*. arXiv preprint arXiv:1902.11097, 2019. Disponível em: <https://arxiv.org/pdf/1902.11097.pdf>. Acesso em 30 de agosto de 2020

WORLD INTELLECTUAL PROPERTY ORGANIZATION. *WIPO Technology Trends 2019: Artificial Intelligence*. Geneva, 2019.

YE, Guixin; TANG, Zhanyong; FANG, Dingyi; ZHU, Zhanxing; FENG, Yansong; XU, Pengfei; CHEN, Xiaojiang; WANG, Zheng. *Yet Another Text Captcha Solver: A Generative Adversarial Network Based Approach*. In: Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security (CCS '18). New York: Association for Computing Machinery, 2018. p. 332–348. Disponível em: <https://dl.acm.org/doi/10.1145/3243734.3243754>. Acesso em 16 de março de 2020.

### Outras fontes:

ACCESS NOW. *The Toronto Declaration: Protecting the right to equality and non-discrimination in machine learning systems*. Disponível em: [https://www.accessnow.org/cms/assets/uploads/2018/08/The-Toronto-Declaration\\_ENG\\_08-2018.pdf](https://www.accessnow.org/cms/assets/uploads/2018/08/The-Toronto-Declaration_ENG_08-2018.pdf). Acesso em 31 de outubro de 2020

CARLONI, Giovanna; BIONI, Bruno Ricardo; SOUZA, Carlos Affonso. *Vigilância digital contra o Covid-19: um mal necessário?* Videoconferência disponível em: <https://youtu.be/oBjQAfaf-xc>. Acesso em 8 de abril de 2020.

AGÊNCIA BRASIL. *Agência Brasil explica: o que é a tecnologia 5G*. Disponível em: <https://agenciabrasil.etc.com.br/geral/noticia/2020-03/agencia-brasil-explica-o-que-e-tecnologia-5g>. Acesso em 3 de junho de 2020.

BRAUNER, Daniela F. (via portal Medium). *A semântica de dados é um futuro possível para o aprendizado de máquina?* Disponível em: <https://medium.com/@danibrauner/a-sem%C3%A2ntica-de-dados-%C3%A9-um-futuro-poss%C3%ADvel-para-o-aprendizado-de-m%C3%A1quina-9344661b5db7>. Acesso em 31 de outubro de 2020.

CGS. *CGS Survey Reveals ‘Sustainability’ Is Driving Demand and Customer Loyalty*. Disponível em: <https://www.cgsinc.com/en/infographics/CGS-Survey-Reveals-Sustainability-Is-Driving-Demand-and-Customer-Loyalty>. Acesso em 30 de outubro de 2020.

COMPUTER WEEKLY. *IT chiefs recognise the risks of artificial intelligence bias*. Disponível em: <https://www.computerweekly.com/news/252474408/IT-chiefs-recognise-the-risks-of-artificial-intelligence-bias>. Acesso em 30 de outubro de 2020.

CONJUR. *As lacunas da Lei Geral de Proteção de Dados em casos como o do FaceApp*. Disponível em: <https://www.conjur.com.br/2019-jul-24/renan-lopergolo-lacunas-lgpd-casos-faceapp>. Acesso em 12 de fevereiro de 2020.

DOTEVERYONE. *Workers' View*. <https://www.doteveryone.org.uk/report/workersview/>. Acesso em 30 de outubro de 2020.

EL PAÍS. *Na verdade, o que [...] é exatamente um algoritmo?* Disponível em: [https://brasil.elpais.com/brasil/2018/03/30/tecnologia/1522424604\\_741609.html](https://brasil.elpais.com/brasil/2018/03/30/tecnologia/1522424604_741609.html). Acesso em 7 de junho de 2020.

ENTREPRENEUR. *How Startups Develop and Deploy Matching Algorithms*. Disponível em: <https://www.entrepreneur.com/article/338442>. Acesso em 12 de novembro de 2020.

FOREIGN AFFAIRS. *The Lawless Realm Countering the Real Cyberthreat*. Disponível em: <https://www.foreignaffairs.com/articles/world/2020-10-13/lawless-realm>. Acesso em 31 de outubro de 2020.

HARVARD BUSINESS REVIEW. *Want Less-Biased Decisions? Use Algorithms*. Disponível em: <https://hbr.org/2018/07/want-less-biased-decisions-use-algorithms>. Acesso em 25 de outubro de 2020

INSTITUTO HUMANITAS. *“O big data apresenta uma multimetodologia”*. Entrevista com Walter Sosa Escudero. Disponível em: <http://www.ihu.unisinos.br/78-noticias/593456-o-big-data-apresenta-uma-multimetodologia-entrevista-com-walter-sosa-escudero>. Acesso em 22 de fevereiro de 2020.

MIC. *Crime-prediction tool PredPol amplifies racially biased policing, study shows*. Disponível em: <https://www.mic.com/articles/156286/crime-prediction-tool-pred-pol-only-amplifies-racially-biased-policing-study-shows>. Acesso em 25 de outubro de 2020.

NEXO. *O que está em jogo quando você dá seu CPF na hora da compra*. Disponível em: [https://www.nexojornal.com.br/expresso/2020/02/15/O-que-est%C3%A1-em-jogo-quando-voc%C3%AA-d%C3%A1-seu-CPF-na-hora-da-compra?utm\\_medium=Social&utm\\_campaign=Echobox&utm\\_source=Twitter#Echobox=1582314608](https://www.nexojornal.com.br/expresso/2020/02/15/O-que-est%C3%A1-em-jogo-quando-voc%C3%AA-d%C3%A1-seu-CPF-na-hora-da-compra?utm_medium=Social&utm_campaign=Echobox&utm_source=Twitter#Echobox=1582314608). Acesso em 22 de fevereiro de 2020.

\_\_\_\_\_. *A parcialidade dos algoritmos*. Disponível em: <https://www.nexojornal.com.br/externo/2019/11/24/A-parcialidade-dos-algoritmos>. Acesso em 25 de outubro de 2020.

PHENOMENAL WORLD. *The Long History of Algorithmic Fairness*. Disponível em: <https://phenomenalworld.org/analysis/long-history-algorithmic-fairness>. Acesso em 25 de outubro de 2020.

PR NEWSWIRE. *New Workplace Survey Finds Nearly 80% of Employers Aren't Worried About Unethical Use of AI -- But Maybe They Should Be*. Disponível em: <https://www.prnewswire.com/news-releases/new-workplace-survey-finds-nearly-80-of-employers-arent-worried-about-unethical-use-of-ai--but-maybe-they-should-be-300911214.html>. Acesso em 30 de outubro de 2020.

PROPUBLICA. *Machine Bias - There's software used across the country to predict future criminals. And it's biased against blacks*. Disponível em:

<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.

Acesso em 25 de outubro de 2020.

QUARTZ. *Millions of Facebook users have no idea they're using the internet*. Disponível em: <https://qz.com/333313/millions-of-facebook-users-have-no-idea-theyre-using-the-internet/>.

Acesso em 30 de outubro de 2020.

REINO UNIDO. *Windrush Lessons Learned Review*. Disponível em: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/876336/6.5577\\_HO\\_Windrush\\_Lessons\\_Learned\\_Review\\_LoResFinal.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/876336/6.5577_HO_Windrush_Lessons_Learned_Review_LoResFinal.pdf). Acesso em 12 de novembro de 2020.

REUTERS. *U.N. investigators cite Facebook role in Myanmar crisis*. Disponível em: <https://www.reuters.com/article/us-myanmar-rohingya-facebook-idUSKCN1GO2PN>. Acesso em 30 de outubro de 2020.

Site oficial do Instituto Charles Babbage. Disponível em: <http://www.cbi.umn.edu/about/babbage.html>. Acesso em 3 de julho de 2020.

Site oficial do projeto Algorithmic Justice League. Disponível em: <https://www.ajlunited.org/>. Acesso em 12 de novembro de 2020.

Site oficial do projeto Coronamap. Disponível em: <https://coronamap.site/>. Acesso em 16 de março de 2020.

Site oficial do projeto GPT-3. Disponível em: <https://openai.com/blog/openai-api/>. Acesso em 23 de agosto de 2020.

Site oficial do projeto Iamus. Detalhes disponíveis em: <http://www.geb.uma.es/melomics/melomics.html>. Acesso em 4 de julho de 2020.

THE ATLANTIC. *The Curious Connection Between Apps for Gay Men and Sex Offenders*. Disponível em: <https://www.theatlantic.com/technology/archive/2011/04/the-curious-connection-between-apps-for-gay-men-and-sex-offenders/237340>; Acesso em 25 de outubro de 2020.

THE GUARDIAN. *Home Office to scrap 'racist algorithm' for UK visa applicants*. Disponível em: <https://www.theguardian.com/uk-news/2020/aug/04/home-office-to-scrap-racist-algorithm-for-uk-visa-applicants>. Acesso em 24 de outubro de 2020.

\_\_\_\_\_. *I asked Tinder for my data. It sent me 800 pages of my deepest, darkest secrets*. Disponível em: <https://www.theguardian.com/technology/2017/sep/26/tinder-personal-data-dating-app-messages-hacked-sold>. Acesso em 22 de outubro de 2020.

\_\_\_\_\_. *US senator: Huawei 5G is like Soviets building west's submarines*. Disponível em: <https://www.theguardian.com/technology/2020/jun/02/us-senator-huawei-5g-is-like-soviets-building-wests-submarines>>. Acesso em 3 de junho de 2020.

THE NEW YORKER. *What Can America Learn from Europe About Regulating Big Tech?* Disponível em: <https://www.newyorker.com/tech/annals-of-technology/what-can-america>

[learn-from-europe-about-regulating-big-tech?utm\\_source=twitter&utm\\_medium=social&utm\\_campaign=onsite-share&utm\\_brand=the-new-yorker&utm\\_social-type=earned](#). Acesso em 31 de outubro de 2020.

THE NEW YORK TIMES. *In Coronavirus Fight, China Gives Citizens a Color Code, With Red Flags*. Disponível em: <https://www.nytimes.com/2020/03/01/business/china-coronavirus-surveillance.html>. Acesso em 16 de março de 2020.

THE VERGE. *Google and Microsoft warn investors that bad AI could harm their brand*. Disponível em: <https://www.theverge.com/2019/2/11/18220050/google-microsoft-ai-brand-damage-investors-10-k-filing>. Acesso em 30 de outubro de 2020

UNIÃO EUROPEIA. *Ethics guidelines for trustworthy AI*. Disponível em: <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>. Acesso em 30 de outubro de 2020.

UNIVERSIDADE DE HARVARD. Berkman Klein Center for Internet & Society. *Ethics and Governance of AI*. Disponível em: <https://cyber.harvard.edu/topics/ethics-and-governance-ai>. Acesso em 30 de outubro de 2020.