

ESCOLA DE DIREITO
CURSO DE DIREITO

MARCELLO WETTER FLORIANI

**A PROBLEMÁTICA DA DISCRIMINAÇÃO ALGORÍTMICA: UMA ANÁLISE SOBRE O
QUADRO DE INTELIGÊNCIA ARTIFICIAL NA ESFERA JURÍDICA E SOCIAL**

Porto Alegre
2022

GRADUAÇÃO



Pontifícia Universidade Católica
do Rio Grande do Sul

A PROBLEMÁTICA DA DISCRIMINAÇÃO ALGORÍTMICA: UMA ANÁLISE SOBRE O QUADRO DE INTELIGÊNCIA ARTIFICIAL NA ESFERA JURÍDICA E SOCIAL

Marcello Wetter Floriani¹
Regina Linden Ruaro²

RESUMO

O presente artigo possui como finalidade analisar o fenômeno da discriminação em algoritmos de inteligência artificial e seus efeitos no ordenamento jurídico. Acerca desse tema, por intermédio de uma pesquisa básica exploratória, qualitativa sobre o assunto e de uma pesquisa bibliográfica, objetiva-se, pelo método dedutivo, abordar a problemática, clarificar o entendimento a respeito do assunto e desenvolver a argumentação crítica ao método de atribuição de valores de algoritmos de inteligência artificial e suas consequências. Pretendeu-se a desenvolver o conceito de inteligência artificial, de aprendizado de máquina, de discriminação, de algoritmo e de discriminação algorítmica e enquadrá-los no ordenamento jurídico. Assim como providenciar uma reflexão crítica a respeito da evolução da inteligência artificial e seu impacto na esfera jurídica e social.

Palavras-chave: Inteligência artificial; Discriminação algorítmica; Identidade individual; Lei geral de proteção de dados.

1 INTRODUÇÃO

Nas últimas décadas, teve-se uma aceleração descomunal no desenvolvimento e na adoção de novas tecnologias, embora muitas lacunas ainda existam nesse sentido em diferentes partes do mundo, especialmente nos países menos desenvolvidos. Essa rápida mudança tecnológica está afetando todas as áreas da sociedade, inclusive na aplicação do Direito. Assim esse fenômeno apresenta novos desafios para o operador do Direito. A rápida mudança tecnológica pode ultrapassar a capacidade dos governos e da sociedade de se adaptarem às mudanças que as novas tecnologias trazem, pois podem alterar os mercados de trabalho, perpetuar as desigualdades e levantar questões éticas.

Nesse contexto, o avanço da tecnologia trouxe consigo um aumento inesperado de dados, fazendo com que medidas para analisar e controlar esses dados fossem tomadas. Surge então o *Big Data*, uma ferramenta tecnológica designada a analisar grandes quantidades de dados e revelar padrões, tendências e associações. Essa revolução de dados incentivou a

¹ Graduando do Curso de Direito pela Pontifícia Universidade Católica do Rio Grande do Sul.

E-mail: Marcello.floriani@edu.pucrs.br

² Orientadora, Doutora em Direito pela Universidad Complutense de Madrid (1993) e possui pós-doutorado pela Universidade San Pablo CEU (2006-2008). É procuradora federal da Advocacia-Geral da União aposentada. Atualmente é professora titular da Pontifícia Universidade Católica do Rio Grande do Sul e membro da Comissão Coordenadora do Programa de Pós-graduação em Direito do Estado da Faculdade de Direito. Compõe o Grupo Internacional de Pesquisa "Protección de Datos, Transparencia y Acceso a la Información". Decana Associada da Escola de Direito (2018/2021), Membro do Comitê Gestor do Biobanco da PUCRS, Membro Honorário do Instituto Internacional de Estudos de Direito do Estado - IEDE. Lidera o Grupo de Pesquisa cadastrado no CNPq: Proteção de Dados Pessoais e Direito Fundamental de Acesso à Informação no Estado Democrático de Direito na linha de Tecnologia, Direito e Inovação.

E-mail: ruaro@pucrs.br

produção de sistemas e mecanismos que automatizassem a análise desses grandes bancos de dados. Desse modo surgiu a Inteligência Artificial (IA).

O campo da Inteligência Artificial é vasto, combina a ciência da computação com um conjunto de dados robusto para permitir a resolução de problemas. Dentro desse campo existe o subcampo de *Machine Learning* (ML), também conhecido como aprendizado de máquina, que faz com que os sistemas utilizados pela inteligência artificial sejam capazes de aprender e adaptar-se sem seguir instruções explícitas, usando algoritmos e modelos estatísticos para analisar e tirar inferências de padrões em dados.

Essa manipulação autônoma de dados gera preocupações significativas a respeito dos resultados advindos desses sistemas. Pois, se alimentado de forma equivocada, distorcida da realidade ou viciada, esses dados fornecidos para seu funcionamento farão com que sejam replicados comportamentos e conclusões equivocadas. Isso faz com que exista a possibilidade de algoritmos perpetuarem ou criarem discriminações no seu modelo de resolução de problemas.

Dessa forma, o presente artigo visa a analisar os impactos produzidos por esses sistemas e levantar questionamentos a respeito da forma pela qual os valores adotados por esses mecanismos são condizentes com o ordenamento jurídico, com os princípios fundamentais do Direito, bem como aos efeitos que produzirão na sociedade.

2 INTELIGÊNCIA ARTIFICIAL

A abordagem da presente tese envolve alguns conceitos principais (como inteligência artificial e aprendizado de máquina) que serão previamente discorridos para que se possa ter o entendimento completo do assunto. Os temas aqui citados não terão profunda qualidade técnica, somente serão trazidos para que se tenha clareza na abordagem do assunto principal.

2.1 EVOLUÇÃO HISTÓRICA E CONCEITO DE INTELIGÊNCIA ARTIFICIAL

Na primeira metade do século XX, o interesse científico familiarizava o mundo com a concepção de robôs inteligentes. Esse crescente interesse pelo desenvolvimento científico fez com que fosse criada uma geração de cientistas, matemáticos e filósofos que possuíssem o conceito de inteligência artificial assimilado em suas mentes. Uma dessas pessoas foi o matemático britânico Alan Turing, lembrado como o “pai da ciência da computação”, que explorou essa possibilidade de inteligência artificial. Em 1950, Turing deduziu que se os seres humanos usam as informações disponíveis, bem como a razão, para resolver os problemas e tomar decisões, os computadores poderiam fazer o mesmo. Essa foi a estrutura lógica de seu artigo, intitulado *Computing Machinery and Intelligence*, no qual discutiu como construir máquinas inteligentes e como testar sua “capacidade intelectual”. Nesse artigo, Turing oferece uma pergunta: “*Can machines think?*”³, que instigou o conhecimento que temos hoje de ciência da computação. Ainda no artigo, desenvolveu o que conhecemos hoje por “teste de Turing”, onde um interrogador humano tentaria distinguir entre uma resposta de um computador e a de um ser humano.

Pouco tempo depois, em 1956, o conceito de inteligência artificial teria seu início pelo programa chamado *Logic Theorist*, produzido por Allen Newell, Herbert Simon e J. C. Shaw. É considerado por muitos como o primeiro programa de inteligência artificial. O programa foi apresentado em *Dartmouth Summer Research Project on Artificial Intelligence*, organizado por

³ TURING, A. M. Computing machinery and intelligence. *Mind*, LIX, n. 236, p. 433 – 460, outubro 1950. Disponível em: <https://bit.ly/2PCToam>. Acesso em: 17/04/2022.

John McCarthy e Marvin Minsky. Nessa histórica conferência, McCarthy, imaginando um grande esforço colaborativo, reuniu os principais pesquisadores de diversas áreas para uma discussão sobre inteligência artificial, termo que ele cunhou no próprio evento e no qual ficou definido como “A ciência e engenharia de fazer máquinas inteligentes”⁴. Esse evento impulsionou os próximos vinte anos de pesquisa em inteligência artificial.

Até 1970, a inteligência artificial teve um crescimento rápido. Os computadores evoluíram, podiam armazenar mais informações e se tornaram mais rápidos e baratos. Os algoritmos de aprendizado de máquina melhoraram e diversas instituições privadas e governamentais investiram no futuro da inteligência artificial. O governo estava particularmente interessado na área de processamento da linguagem, estava procurando por uma máquina que pudesse transcrever e traduzir a linguagem falada, e, por consequência, essa foi a área que mais cresceu nesse período, onde tiveram invenções como ELIZA por Joseph Weizenbaum, que foi o primeiro software para simulação de diálogos. Entretanto, apesar do crescente incentivo para a IA, os projetos não conseguiam se efetivar devido às condições tecnológicas da época, fazendo com que houvesse diversos cortes, redução de projetos ativos, e redução significativa de investimento na área.

Em 1990, apesar da ausência de financiamento do governo e interesse público, a inteligência artificial prosperou e começou a bater os marcos de desenvolvimento. Em 1997, o então atual campeão de xadrez, Gary Kasparov, foi derrotado pelo software “*Deep Blue*” da IBM, um programa de computador de jogar xadrez. Esse foi um marco histórico para o desenvolvimento da inteligência artificial. No mesmo ano, um software de reconhecimento de voz foi implementado ao Windows, e o movimento de avanço tecnológico continua até hoje.

Atualmente, o problema de armazenamento de dados de antigamente não é mais relevante, estamos vivendo na era do *Big Data*, onde temos a capacidade de coletar enormes somas de informações e fazer com que elas sejam processadas por algoritmos de inteligência artificial por diferentes técnicas, como *Machine Learning*.

2.2 MACHINE LEARNING

Machine Learning, ou aprendizado de máquina, é uma coleção de métodos que permitem que os computadores automatizem a construção e programação de algoritmos orientados por dados, por meio de uma descoberta sistemática de padrões nos dados alimentados. Em 1959, Arthur Samuel definiu aprendizado de máquina como um “campo de estudo que dá aos computadores a capacidade de aprender sem serem explicitamente programados”⁵.

Com a evolução das tecnologias de computação e comunicação, tornou-se possível utilizar esses algoritmos de aprendizado de máquina para identificar padrões cada vez mais complexos e ocultos nos dados. Além disso, atualmente é possível desenvolver modelos que se adaptam automaticamente à base de dados maiores e mais complexas, e ajudar a estimar o impacto de decisões de diversos cenários a respeito do assunto em tempo real.

O algoritmo de aprendizado de máquina é o método pelo qual o sistema de IA realiza sua tarefa, geralmente prevendo os dados de saída (*output*) sem relação aos dados de entrada (*input*) inseridos. Os dois principais processos algoritmos de aprendizado de máquina são classificação e regressão. Os algoritmos de aprendizado de máquina são amplamente

⁴ MCCARTHY, J. “the science and engineering of making intelligent machines.”. 1955.

⁵ Tradução livre. SAMUEL, A. L. Some Studies in Machine Learning Using the Game of Checkers. IBM Journal of Research and Development, v. 3, n. 3, p. 210 – 229, 1959. Disponível em: <http://dx.doi.org/10.1147/rd.33.0210>. Acesso em: 18/04/2022.

categorizados como supervisionados ou não supervisionados. Os algoritmos de aprendizado supervisionado têm dados de entrada e dados de saída desejados fornecidos por meio de rotulagem, enquanto os algoritmos não supervisionados trabalham com dados que não são classificados nem rotulados. Na medida em que os sistemas evoluem, a complexidade abordada aumenta e faz com que fatores não previamente esperados ocorram.

3 DISCRIMINAÇÃO NOS ALGORITMOS DE INTELIGÊNCIA ARTIFICIAL

Conforme os algoritmos expandem sua capacidade de organizar a sociedade, a política, as instituições e o comportamento, a população se preocupa com as maneiras pelas quais a produção distorcida e a manipulação de dados podem impactar as relações sociais. Para se ter uma visão clara do assunto, primeiro abordar-se-ão os motivos que originaram a pauta do título e o contexto histórico em que o assunto se desenvolveu.

3.1 EVOLUÇÃO HISTÓRICA DA DISCRIMINAÇÃO NOS ALGORITMOS DE INTELIGÊNCIA ARTIFICIAL

Vieses (tendência geral ou determinada por forças externas)⁶ e discriminação, em um primeiro momento, podem parecer questões exclusivamente humanas. Programas de software e sistemas de algoritmos, em teoria, deveriam ser imunes a esses conceitos, afinal, eles aplicam as mesmas regras baseadas em sua hierarquia de valores programados. Porém, na realidade, os efeitos que alguns algoritmos produzem estão sendo considerados discriminatórios, mesmo que sejam blindados do subjetivismo humano.

Embora seja um assunto relevante atualmente, a tomada de decisão por algoritmos já era motivo de discussão até mesmo antigamente. Em 1976, Joseph Weizenbaum, defendia a ideia de que, mesmo sendo possível a utilização de inteligência artificial na época, nunca se deveria deixá-la tomar decisões importantes, pois sempre lhes faltaria qualidades humanas, como compaixão e sabedoria⁷. Ideia que fomentou o que atualmente é a discussão de “consciência” da inteligência artificial. Na época, Weizenbaum fez a distinção entre escolher e decidir, sendo decidir algo programável, que um computador poderia avaliar matematicamente e chegar a uma decisão. Porém, segundo ele, escolha seria fruto de julgamento, e não cálculo matemático, e o julgamento humano é capaz de incluir fatores não matemáticos, como a emoção e o subjetivismo. Para o autor, a capacidade de escolher é o que nos torna fundamentalmente humanos. Weizenbaum escreveu que os programas são uma sequência de regras criadas por humanos para um computador seguir.

Ao seguir essas regras de forma consistente, esses programas impõem uma maneira específica de resolver problemas. As regras que um computador segue são baseadas nas suposições de um programador de computador sobre como esses problemas podem ser resolvidos. Isso significa que o código pode incorporar a imaginação do programador de como o mundo funciona, incluindo seus preconceitos e expectativas.

Essa é a ideia central de discriminação algorítmica, quando se fala do assunto, apontar que o processo de sua criação não está livre de incorporar as características do programador, e, portanto, pode virar um meio de perpetuar em programas os vieses e preconceitos de seu criador. A discussão a respeito do assunto se torna difícil, pois as decisões tomadas por um

⁶ VIÉS. *In*: GOOGLE, Oxford Languages. 2022.

⁷ WEIZENBAUM, J. *Computer Power and Human Reason: From Judgment to Calculation*. [S.l.]: W. H. Freeman and Company, 1976.

programa sem percepções humanas, como preconceito e viés, tendem a ter mais credibilidade que as decisões tomadas por humanos.

3.2 AUTORIDADE DAS DECISÕES DE ALGORITMOS DE INTELIGENCIA ARTIFICIAL

Embora algoritmos bem projetados frequentemente determinem resultados que são semelhantes ou até mais justos que as decisões de seres humanos, casos de viés ainda ocorrem e são difíceis de prever e analisar. A complexidade necessária para analisar casos de viés algorítmico cresceu junto com a complexidade dos programas a serem analisados. Alguns sistemas contam com diversos programadores, que podem ter suas decisões infiltradas entre as diversas linhas de códigos de um programa. Mas, devido a essa tamanha complexidade para que se tenha a existência desses programas, as decisões de algorítmicos podem ser vistas como mais autoritárias do que as decisões dos seres humanos. De acordo com O’neil (2016, p.5):

As aplicações matemáticas fomentando a economia dos dados eram baseadas em escolhas feitas por seres humanos falíveis. Algumas dessas escolhas sem dúvida eram feitas com as melhores das intenções. Mesmo assim, muitos desses modelos programavam preconceitos, equívocos e vieses humanos nos sistemas de software que cada vez mais geriam nossas vidas. Como deuses, esses modelos matemáticos eram opacos, seus mecanismos invisíveis a todos, exceto os altos sacerdotes de seus domínios: os matemáticos e cientistas da computação. Suas decisões, mesmo quando erradas ou danosas, estavam para além de qualquer contestação.⁸

O’neil argumenta que, por causa de sua conveniência e autoridade, os algoritmos são teorizados como um meio de delegar a responsabilidade dos humanos. No momento em que o agente que está tomando a decisão não é um ser humano, por algum motivo, o resultado se torna mais credível, mesmo sendo programados por seres humanos. Essa falácia lógica é conhecida como “argumento de autoridade”. Embora seja uma proposição válida em termos de atribuir credibilidade ao argumento, passa a ser uma falácia lógica quando se utiliza do argumento em si para provar a veracidade da proposição.

Argumentum ad verecundiam ou argumento de autoridade, também chamado de apelar à autoridade, é uma forma de argumento em que a opinião de uma autoridade sobre um tópico é usada como fundamento para evidenciar a credibilidade de um argumento lógico. É uma falácia lógica, pois a fonte está sendo usada para justificar a aceitação de uma conclusão. Essa falácia pode ser difícil de evitar porque geralmente temos boas razões para acreditar em autoridades ou figuras de especialistas. No entanto, é importante perceber que a validade de um argumento não tem a ver com a pessoa que faz a alegação, e, sim, nas evidências que ela apresenta. Recentemente, Ben Shapiro, em uma de suas palestras, fez alusão ao argumento de autoridade. Segundo ele:

“[...] Eu não acho que colocar um PhD da Ferris State ao lado do seu nome faz de você um especialista em todas as coisas no campo em que você se propõe a falar. O fato é que ou meu argumento é bom ou meu argumento é ruim. Citar minhas credenciais é uma maneira muito ruim de argumentar. Na verdade, é o equivalente [...] é o equivalente a uma pessoa religiosa citando a Bíblia como argumento. Eu sou uma pessoa muito religiosa. [...] Eu sou uma pessoa religiosa. Eu nunca cito a Bíblia. A razão pela qual eu não cito a Bíblia é porque isso é um argumento de autoridade. Você pode não acreditar na autoridade a que estou citando. Eu não acredito na

⁸ O’NEIL, C. Algoritmos de Destruição em Massa. [S.l.]: Editora Rua do Sabão, 2021.

autoridade à qual você está citando, então você vai ter que me argumentar sobre o que estou dizendo que está errado, em vez de apenas dizer que não tenho o diploma devidamente enumerado da instituição de sua preferência”.⁹

Essa dificuldade em ter uma discussão a respeito do assunto e a dificuldade do assunto em si postergaram o diálogo a respeito do tema. Entretanto, o aumento das consequências advindas da discriminação algorítmica trouxe à tona a necessidade de construir-se uma inserção mais moderna sobre o assunto.

3.3 CONCEITO DE DISCRIMINAÇÃO ALGORÍTMICA

Tendo abordado o desenvolvimento histórico do tema, agora é necessário definir o que se entende pela expressão discriminação algorítmica. Primeiramente, há de se clarificar o que se entende por discriminação. Thomas Sowell (2019, p.29) aborda a classificação de duas formas:

Diz-se que algumas pessoas possuem gosto discriminativo quando são especialmente boas em detectar diferenças de qualidade na hora de escolher vinhos, pinturas ou outros bens e serviços. Mas a palavra também é usada, no sentido quase oposto, para se referir a diferenças arbitrárias de comportamento em relação às pessoas com base em suas identidades grupais, quaisquer que sejam suas qualidades reais como indivíduo.¹⁰

Da mesma forma, Adilson José Moreira (2017, p.27), reconhece, também, a pluralidade de significados e define a palavra:

A palavra discriminação possui uma pluralidade de significados, embora tenha adquirido um sentido bem específico no mundo atual. Ela designa, por um lado, a ação de classificar objetos a partir de um determinado critério. [...] Hoje, o termo discriminar tem conotações claramente negativas, pois sugere que alguém foi tratado de forma arbitrária. Os dois sentidos da palavra estão presentes no vocabulário jurídico. [...] O vocábulo discriminar significa aqui categorizar pessoas a partir de uma característica ou situação jurídica para atribuir a elas alguma consequência. Contudo a palavra discriminação tem outro significado no mundo do Direito: ela indicia que uma pessoa impõe a outra um tratamento desvantajoso a partir de um julgamento moral negativo.¹¹

Ante o exposto, o autor classifica discriminação como a ação de classificar objetos a partir de um determinado critério, veja bem, discriminação só ocorre de fato quando há o ato de agir baseado em um determinado critério. Assim como no outro significado que o autor apresenta, a manifestação do julgamento moral negativo através da ação de impor algo a outro. Muitas vezes, esse conceito de discriminação é confundido com o conceito de preconceito.

Preconceito carece do ato de agir, segundo dicionário¹²: “qualquer opinião ou sentimento concebido sem exame crítico.”. Há de se concluir que, não existe a ideia de ação, e, sim, um simples sentimento em relação a algo. De acordo com Gordon Allport (1954, p.9), o preconceito pode ser definido como: “uma antipatia baseada em uma generalização defeituosa

⁹ Tradução livre. SHAPIRO, B. WATCH: Ben Shapiro Breaks Down The ‘Argument From Authority’. Daily Wire, 2017. Disponível em: <https://www.dailywire.com/news/watch-ben-shapiro-breaks-down-argument-authority-hank-berrien>. Acesso em: 28/04/2022.

¹⁰ SOWELL, T. Discriminação e disparidades. [S.l.]: Record, 2019.

¹¹ MOREIRA, A. J. O que é discriminação? [S.l.]: Casa do Direito, 2017.

¹² PRECONCEITO. In: GOOGLE, Oxford Languages. 2022

e inflexível. Pode ser dirigido a um grupo como um todo, ou a um indivíduo por ele ser um membro desse grupo.”. Adicionalmente, Rupert Brown (2018, p.8) comenta ao conceito:

[...] o preconceito será considerado como qualquer um ou todos os seguintes: a manutenção de atitudes sociais ou crenças cognitivas depreciativas, a expressão de afeto negativo ou a exibição de comportamento hostil ou discriminatório em relação a membros de um grupo por causa de ser membro desse grupo.¹³

Posto isso, entende-se que a diferença entre os dois é que a discriminação é um ato de tratamento injusto dirigido a um indivíduo ou a um grupo e o preconceito é uma atitude ou prejulgamento, geralmente de forma negativa. A discriminação pode manifestar-se de duas formas, a discriminação positiva e a discriminação negativa. A discriminação positiva, também chamada de *Affirmative Action*, é o esforço ativo para apurar o resultado dos sistemas de escolha para membros de grupos que foram submetidos a tratamentos de distinção. Diferente de discriminação negativa, que faz alusão ao conceito clássico de discriminação, conforme Adilson José Moreira (2017, p.30):

A discriminação negativa designa um tratamento que viola o princípio segundo o qual todos os membros de uma comunidade política devem ser igualmente respeitados. Ela acontece quando um agente público ou privado trata uma pessoa de forma arbitrária, o que é frequentemente motivado por estigmas culturais. [...] Vemos então que a discriminação negativa está baseada em uma motivação ilegítima [...].¹⁴

Embora nítida a definição, a discriminação nem sempre ocorre de maneira clara e evidente. Essa pode ocorrer de maneira direta ou indireta, portanto, faz-se necessário a abordagem desses conceitos.

A discriminação direta refere-se à ação de tratar uma pessoa de forma diferente (muitas vezes desfavorável) devido a uma característica física ou mental específica: como idade, sexo ou deficiência. Ademais, a discriminação direta é aquela que contém em si o animus, a intenção de discriminar. O sujeito é discriminado intencionalmente, tratado de forma desigual, em função de uma característica, individual ou de grupo, que lhe é peculiar, tal como sua condição física (RIBEIRO, 2018)¹⁵.

No tocante a discriminação indireta, a diferença de tratamento aparece de forma indireta, dissimulada, desprovida de fator de intencionalidade, cujos efeitos advêm de práticas ou políticas aparentemente neutras, mas que redundam em atos discriminatórios (RIBEIRO, 2018). Conhecida, também, como teoria do impacto desproporcional, a discriminação indireta resultou em uma teoria judicial desenvolvida nos Estados Unidos que permite desafios ao emprego ou práticas educacionais que não são discriminatórias em seu teor, mas têm um efeito desproporcionalmente negativo sobre membros de grupos legalmente protegidos. De acordo com Joaquim Barbosa (2001, n.p.):

Toda e qualquer prática empresarial, política governamental ou semi-governamental, de cunho legislativo ou administrativo, ainda que não provida de intenção discriminatória no momento de sua concepção, deve ser condenada por violação do princípio constitucional da igualdade material se, em consequência de sua aplicação,

¹³ Tradução livre. BROWN, R. Prejudice: Its Social Psychology. [S.l.]: Wiley-Blackwell, 2018.

¹⁴ MOREIRA, A. J. O que é discriminação? [S.l.]: Casa do Direito, 2017.

¹⁵ RIBEIRO, V. C. Todos os tipos de discriminação. 2018. Disponível em: <https://valeriacy.jusbrasil.com.br/artigos/583475306/todos-os-tipos-de-discriminacao>. Acesso em: 03/05/2022.

resultarem efeitos nocivos de incidência especialmente desproporcional sobre certas categorias de pessoas¹⁶

Ainda, a falta de acessibilidade adequada para pessoas com deficiência física na atual construção de edifícios públicos e privados pode ser caracterizada como uma forma de discriminação indireta. Por não prover meios adequados de acesso a essas pessoas, acaba criando uma dificuldade de acesso desproporcional.

Feito o delineamento a respeito do conceito de discriminação, passa-se agora a discorrer sobre o que é um algoritmo. Podemos definir algoritmo, de uma maneira simplificada, como uma lista de instruções que determinam como os programas leem, coletam, processam e analisam dados, eles recebem uma sequência de *inputs* e processam uma sequência de *outputs*. Quando tratamos de algoritmos de inteligência artificial, estamos falando de programas complexos, que possuem a capacidade de inferir resultados, assim como de decidir automaticamente a respeito do assunto, conforme sua valoração programada. Isso pode variar desde as recomendações que se pode ver em uma página de pesquisa ou serviço de streaming, aos anúncios dos sites, anúncios de empregos e até o uso de algoritmos para auxílio da aplicação da lei.

Entretanto, nem sempre o resultado de aplicação de algoritmos resulta em benefícios, conforme Sandra Watcher:

[...] algoritmos são cada vez mais usados para nos avaliar, prever nossos comportamentos e preferências e, finalmente, tomar decisões importantes sobre nós. Algoritmos podem inferir nossa orientação sexual, posição política e estado de saúde sem que estejamos cientes. Eles também decidem quais produtos ou feeds de notícias são mostrados para nós, bem como se somos contratados, demitidos ou promovidos, se recebemos um empréstimo ou seguro, entre muitas outras coisas. Os algoritmos fazem isso extraindo inferências de dados altamente diversos e ricos em recursos (como nosso histórico de navegação na web ou interações em redes sociais). Essas inferências são muitas vezes invasivas, contra-intuitivas e não verificáveis, e somos incapazes de prevê-las, compreendê-las ou refutá-las. No entanto, essas inferências moldam nossas identidades e reputações e orientam nossos caminhos na vida. Assim, as decisões baseadas em dados criam novas oportunidades para a tomada de decisões discriminatórias e tendenciosas.¹⁷

As inferências resultantes dos algoritmos muitas vezes causam danos e possuem a lógica de suas decisões enviesadas, onde acabam criando erros sistemáticos repetitivos que criam resultados injustos. Essa é a concepção central de discriminação algorítmica. Chama-se de discriminação algorítmica os erros sistemáticos e repetitivos que criam resultados injustos, como privilegiar um grupo arbitrário de usuários em detrimento de outros. Adicionalmente, Nima Kordzaech e Maryam Ghasemaghahi comentam:

[...]Assim, processos algorítmicos que são usados para automatizar ou auxiliar a tomada de decisões sobre pessoas podem produzir resultados discriminatórios que violam as normas de justiça e igualdade e que impactam negativamente determinadas pessoas ou comunidades no local de trabalho ou na sociedade. Esse fenômeno é chamado de viés algorítmico e ocorre quando as saídas de um algoritmo beneficiam

¹⁶ GOMES, Joaquim Barbosa. Ação Afirmativa e Princípio Constitucional da Igualdade. Rio de Janeiro: Renovar, 2001

¹⁷ Tradução livre. WATCHER, S. Making algorithms accountable and explainable: the need for a legal framework. Disponível em: <https://www.research.ox.ac.uk/article/2018-10-15-making-algorithms-accountable-and-explainable-the-need-for-a-legal-framework>.

ou desfavorecem certos indivíduos ou grupos mais do que outros sem uma razão justificada para tais impactos desiguais.¹⁸

Essa abordagem é mais popular em relação ao assunto, porém não é a única. Tratar apenas de vieses e consequências pode ser a maneira mais prática de desenvolver o tema, entretanto, o assunto permite uma reflexão mais basilar a respeito do método de julgamento.

3.4 DESCONSIDERAÇÃO DA IDENTIDADE INDIVIDUAL POR ALGORITMOS DE INTELIGÊNCIA ARTIFICIAL

Há outro problema fundamental quando debatemos o assunto. O método de julgamento de algoritmos possui uma falha ao utilizarem características de personalidade de grupo. Algoritmos inferem um raciocínio lógico matemático em algo complexo e relativo. Atribuir a valoração de comportamentos à um indivíduo a partir de uma análise de um grupo seletivo de dados pode ser discriminatório. Esse tópico vem encontrando bastante espaço na mídia atualmente, conforme vimos pelo conceito de discriminação positiva, principalmente no ramo da política. Ele vem sendo identificado como *Identity politics*, ou política identitária, que é uma abordagem política em que pessoas de uma determinada raça, religião, gênero, origem social, classe social ou outros fatores de identificação desenvolvem agendas políticas baseadas nessas identidades.

A filosofia de atribuir soluções a um problema de grupo é perigosa pelo simples fato de que todas as pessoas são diferentes entre si, possuem diferentes experiências e realidades. Conglomerar um grupo de pessoas como sendo idênticas e atribuir soluções genéricas ao problema é ilógico. Dito isso, é claro que alguns problemas são os mesmos para alguns indivíduos e as soluções podem ser as mesmas. Destarte, como os indivíduos se diferem biologicamente, as soluções variam da mesma maneira. Assim sendo, não se pode ter como base a identidade de grupo para atribuição de soluções singulares. Jordan Peterson, em seu artigo, aborda o tema de identidade:

Sua identidade não é a roupa que você veste, ou a preferência sexual da moda ou comportamento que você adota e ostenta, ou as causas que impulsionam seu ativismo, ou sua indignação moral com ideias diferentes das suas: corretamente entendida, é um conjunto de compromissos complexos entre o indivíduo e sociedade sobre como o primeiro e o segundo podem se apoiar mutuamente de maneira sustentável e de longo prazo. Não é nada para alterar levemente, pois tal compromisso é muito difícil de alcançar, constituindo a essência da própria civilização, que levou eras para se estabelecer e entender, como deveríamos, que a alternativa à adoção de papéis socialmente aceitáveis é conflito — claro, simples e contínuo, bem como simultaneamente psicológico e social.¹⁹

Consoante o artigo, Peterson aborda o fato da importância da identidade individual, e a constante negociação de regras comportamentais com a sociedade.

O que se entende por identidade individual, às vezes chamada de identidade pessoal, em uma explicação rasa, refere-se a propriedades às quais nutrimos um sentimento original de pessoalidade. A identidade pessoal de alguém, nesse sentido, consiste naquelas características

¹⁸ Tradução livre. KORDZAECH, N.; GHASEMAGHAEI, M. Algorithmic bias: review, synthesis, and future research directions. Disponível em: shorturl.at/sFIM1. Acesso em: 10/05/2022.

¹⁹ Tradução livre. PETERSON, J. On the Psychological and Social Significance of Identity. Disponível em: <https://www.jordanbpeterson.com/blog-posts/on-the-psychological-and-social-significance-of-identity/>. Acesso em: 12/05/2022.

que assume para defini-la como pessoa ou torná-la a pessoa que é, e que a distinguem dos outros.

Quanto à identidade de grupo, a identidade individual é secundária e a identidade grupal torna-se parte protagonista da identidade da pessoa. A ideia é que, participando de relações sociais, uma pessoa pode desenvolver um sentimento de pertencimento e uma identidade de grupo que vai além da pessoa individual.

A dificuldade na programação de algoritmos é entendível, visto que é de extrema dificuldade inferir valores matemáticos a características e ainda mais complexo valorar quais características se deve levar em consideração no sistema. Por conseguinte, a inteligência artificial, em algumas instâncias, acaba por utilizar-se de valores coletivos para atribuir a adequação do indivíduo aos parâmetros atribuídos, o que faz com que a individualidade seja perdida no processo.

3.5 EXEMPLOS DE DISCRIMINAÇÃO ALGORÍTMICA

A noção de discriminação algorítmica não é algo novo. Em 1979, tivemos o que ficou marcado como o primeiro caso de discriminação algorítmica. Dr. Geoffrey Franglen fazia parte do processo de admissão para a faculdade de medicina do hospital de *St George*. No seu esforço de automatizar o processo tedioso de admissão, Geoffrey resolveu criar um algoritmo que poderia assumir este processo de triagem inicial. Ele projetou o algoritmo para imitar os processos que ele e seus colegas avaliadores usavam.

Por volta de 1982, o algoritmo já era usado para automatizar todo processo de admissão do hospital, e era entendido que a automação, por ser feita pelo algoritmo, teria menos problemas. Porém, em 1986, após preocupações com a diversidade das contratações, foi aberta uma investigação pela “Comissão de igualdade racial do Reino Unido”²⁰. Então foi descoberto que o algoritmo incluía regras que atribuíam valor tanto no local de nascimento quanto no nome do solicitante e então os enquadravam em “caucasianos” ou “não-caucasianos” a partir dos dados, sendo descreditado pontos ao enquadrar-se na última categoria citada. Ademais, apenas ter um “nome não europeu” podia custar ao candidato 15 pontos da pontuação atribuída pelo algoritmo. Já as mulheres, perdiam em média 3 pontos somente em razão do sexo. Segundo a Comissão:

Há muito tempo que se suspeita de discriminação na medicina contra as mulheres e outros grupos de minorias, mas agora está provado. A St George's Hospital Medical School foi considerada culpada pela comissão de igualdade racial por praticar discriminação racial e sexual em sua política de admissão.²¹

No caso retratado, vemos um claro exemplo de discriminação direta, em virtude do raciocínio lógico que o algoritmo se fez valer para atribuir candidatos aptos para o cargo ao atribuir liberdade ao algoritmo para definir sua própria hierarquia de valores. Ele fez a conexão de valores completamente irrelevantes ao caso, por intermédio de um simples raciocínio matemático. Há também o caso de viés algorítmico, pelo fato de que se havia poucos dados para o sistema se basear, e, nos dados que havia, não existia uma diversidade. Ao ser

²⁰ Tradução livre. “UK Commission for Racial Equality”

²¹ Tradução livre. “Discrimination in medicine against women and members of ethnic minorities has long been suspected, but it has now been proved. St George’s Hospital Medical School has been found guilty by the Commission for racial Equality of practicing racial and sexual discrimination in its admissions policy.” EQUALITY, C. for R. Medical school admissions Report of a formal investigation into St George’s Hospital Medical School. Disponível em: <https://www.jstor.org/stable/pdf/community.28327674.pdf>. Acesso em: 15/05/2022.

programado, o objetivo era filtrar pessoas aptas para o cargo, e, sendo atribuído somente currículos de pessoas anteriormente aprovadas, deduziu que a identidade de grupo dessas pessoas seria um bom parâmetro para aferir a aptidão em relação ao cargo, como foi abordado no tópico anterior.

Existem também, muitos outros casos de discriminação algorítmica. Sob a mesma ótica, a Amazon, atualmente a quinta empresa mais valiosa do mundo, em 2014, cometeu o mesmo erro. Em seu esforço para revisar currículos de candidatos a emprego, criou um programa para facilitar o trabalho, porém tiveram a aproximação errada quanto ao treinamento de seus algoritmos. Os modelos de aprendizado de máquina foram treinados para avaliar os candidatos, observando padrões nos currículos enviados à empresa durante um período de 10 anos.

Como o algoritmo usou os resultados de suas próprias previsões para melhorar sua precisão, ficou preso em um padrão de sexismo contra candidatas do sexo feminino, exatamente da mesma maneira em que aconteceu em 1982, no hospital St George. Pela maneira em que o algoritmo foi programado, ele buscava reproduzir o padrão de currículos aprovados anteriormente, que, em sua maioria, eram homens. A empresa decidiu deixar de fazer o uso do software algum tempo depois.

Casos como esses exemplificam a necessidade de ter-se cuidados com os dados que são alimentados para esses programas, não se pode esperar uma aferição de parâmetros humanos em uma lógica matemática. O programa simplesmente irá utilizar o que foi alimentado para tentar reproduzir da melhor maneira o resultado. Essa falta de valores humanos é uma consequência completamente previsível, visto que o programa não possui o que se entende atualmente como “consciência”.

Há também casos que apresentam outra problemática. Em 2019, a empresa Apple, considerada hoje a segunda empresa mais valiosa do mundo em termos de valor de mercado, lançou um produto chamado *Apple Card* onde foi produzido um algoritmo para calcular o limite de crédito que seria concedido aos usuários que adquirissem o cartão. Entretanto, pouco tempo após o lançamento do produto, notou-se que surgiram reclamações em redes sociais a respeito do algoritmo, onde dizia-se que homens estariam tendo limites maiores em relação às suas esposas, mesmo possuindo as mesmas condições financeiras e viverem no mesmo endereço, como argumentaria Steve Wozniak, co-fundador da empresa, em um dos seus tweets em resposta a uma reclamação a respeito do assunto:

A mesma coisa aconteceu conosco. Eu tenho 10x o limite de crédito. Não temos contas bancárias ou de cartão de crédito separadas ou quaisquer ativos separados. Difícil chegar a um humano para uma correção embora. É a apenas a natureza da *big tech* em 2019.²²

Essas alegações culminaram em uma investigação pelo “Departamento de Serviços Financeiros do Estado de Nova York”²³. As fundamentações para investigação tinham um motivo parecido com a dos casos anteriores. A investigação buscava saber se a empresa teria influenciado para que o algoritmo desenvolvesse um viés negativo em relação ao sexo feminino. Em outras palavras, mesmo que o viés de gênero fosse puramente não intencional, se estivesse

²² Tradução livre. WOZNIAK, S. The same thing happened to us. I got 10x the credit limit. We have no separate bank or credit card accounts or any separate assets. Hard to get to a human for a correction though. It’s big tech in 2019. 2019. Twitter. Disponível em: https://twitter.com/stevevoz/status/1193330241478901760?ref_src=twsrc%5Etfw%7Ctwcamp%5Etweetembed%7Ctwterm%5E1193330241478901760%7Ctwgr%5E%7Ctwcon%5Es1_&ref_url=https%3A%2F%2Fwww.idropnews.com%2Fnews%2Fno-the-apple-card-does-not-discriminate-against-women%2F154988%2F. Acesso em: 20/05/2022.

²³ Tradução livre. “New York State Department of Financial Services”.

presente nos algoritmos de alguma forma, a empresa poderia ser responsabilizada por tratamento discriminatório.

Mas, o cartão tinha uma peculiaridade. O Apple Card não permitia contas conjuntas. Os cônjuges que desejassem seu próprio cartão eram forçados a abrir uma requisição separada, o que pode destacar um problema que simplesmente não é encontrado na maioria dos outros provedores de cartão de crédito.

O resultado, segundo o Departamento de Serviços Financeiros do Estado de Nova York, negou a existência de tratamento diferenciado:

Conforme detalhado abaixo, a análise exaustiva do Departamento de documentação e dados fornecidos pelo banco e pela Apple, juntamente com inúmeras entrevistas de consumidores que se queixaram de possível discriminação, não produziu evidências de discriminação de impacto deliberada ou díspar, mas mostrou deficiências no atendimento ao cliente e na transparência, que o Banco e a Apple tomaram medidas para remediar. Além disso, este relatório aborda questões iluminadas pela discussão pública das alegações, incluindo equívocos sobre “finanças compartilhadas” do cônjuge e usuários autorizados, como construir histórico de crédito, a necessidade de maior transparência nas decisões de crédito, viés inerente à pontuação de crédito e a riscos e benefícios do uso de inteligência artificial na tomada de decisão de crédito...²⁴

Porém, a investigação fez alusão à transparência no tratamento de dados pelos algoritmos, assunto que tem uma ligação muito forte com a discriminação algorítmica. Pois, se não é possível analisar os dados, é impossível chegar a uma conclusão quanto à discriminação. Para fins de se evitar a simplificação excessiva do assunto, não irei abordar transparência no presente artigo. Essa investigação é fundamental ao tratamento da temática do artigo. Muitas vezes, a discrepância no resultado não implica a discriminação no tratamento dos dados pelo algoritmo. É de extrema importância abordar o assunto com cuidado e calma.

Sob outra ótica, Joy Buolamwini, a estudante negra de pós-graduação do MIT²⁵, estava trabalhando com um software de análise facial quando percebeu um problema: o software não detectava o seu rosto, porém, quando usava uma máscara branca em frente a sua face, o algoritmo conseguia detectá-la. A inspiração para sua pesquisa veio quando ela estava no MIT, criando sua arte *Aspire Mirror*²⁶, que usa reconhecimento facial para refletir outra pessoa que o inspira, em seu rosto. Buolamwini antecipou ter Serena Williams, outra mulher de pele escura, refletida em seu rosto. No entanto, a tecnologia não reconheceu seu rosto. Na palestra que realizou no TED, a pesquisadora explicou que investigava o porquê isso estaria ocorrendo e, consequentemente, a levou a concluir que a exclusão de pessoas que se parecem com ela era resultado de um termo que ela apelidou de *Coded Bias*, que subsequentemente virou um documentário no Netflix.

Esse caso introduz uma nova perspectiva em relação ao tema. Por se tratar de um algoritmo de reconhecimento facial, os padrões com que o algoritmo se utiliza para determinar a identificação de uma face não podem ser o mesmo para todas as raças, pois possuem

²⁴ Tradução livre. “As detailed below, the Department’s exhaustive review of documentation and data provided by the Bank and Apple, along with numerous interviews of consumers who complained of possible discrimination, did not produce evidence of deliberate or disparate impact discrimination but showed deficiencies in customer service and transparency, which the Bank and Apple have since taken steps to remedy. Additionally, this report addresses issues illuminated by the public discussion of the allegations, including misconceptions about spousal “shared finances” and authorized users, how to build credit history, the need for increased transparency in credit decisions, bias inherent in credit scoring, and the risks and benefits of using artificial intelligence in credit decision-making.”

²⁵ Massachusetts Institute of Technology

²⁶ BUOLAMWINI, J. *Aspire Mirror*. Disponível em: <http://www.aspiremirror.com>. Acesso em: 01/06/2022.

diferenças significativas em características faciais. Claro, existem mais semelhanças que diferenças, isso se faz evidente, porém para o algoritmo, há de se entender que existem essas diferenças. A maneira com que a luz reflete do rosto de uma pessoa negra não é a mesma de que de um rosto branco. Assim como a circunferência dos olhos não é mesma entre pessoas asiáticas e não asiáticas. O problema do caso relatado acima, é ter seu banco de dados para o aprendizado de máquina contendo uma maioria esmagadora de pessoas caucasianas, que fez com que os padrões de identificação facial funcionassem corretamente somente em pessoas caucasianas e incorretamente em outras raças. Casos com o mesmo problema são comuns atualmente, se parece que o problema de identificação facial não é tratado com a devida importância.

No mesmo sentido, outras empresas foram alvos de críticas pelo mesmo motivo. Em setembro de 2020, a plataforma de videochamadas Zoom foi alvo de críticas após uma postagem na rede social Twitter, em que o aluno na época, Colin Madland, relatou:

Um membro do corpo docente perguntou como impedir que o Zoom removesse sua cabeça enquanto ele usa um plano de fundo virtual. Sugerimos o plano de fundo usual, boa iluminação etc. Mas não funcionou. Eu estava em uma reunião com ele hoje quando percebi por que isso estava acontecendo.²⁷

Collins observou que mesmo as condições de iluminação e escolha de plano de fundo adequado, o programa ainda deixava de mostrar o rosto de seu professor negro. Relatando um problema que impedia o uso dessa função pelo professor. Enquanto o caso em relação ao zoom não se teve muita repercussão, ao postar sua indagação online com uma foto do ocorrido, percebeu que o foco da imagem na postagem era o seu rosto, ao invés do rosto de seu professor, mesmo postando diversas vezes a imagem, até mesmo invertendo-a. Twitter é uma plataforma online de rede social que possui um algoritmo de destacamento de imagem que tem como objetivo otimizar a visualização de imagens no feed de notícias quando se tem uma imagem muito grande, fazendo com que se tenha o foco em determinadas partes da imagem, muitas vezes o rosto. Porém, Collins, ao fazer sua indagação, fez com que se desencadeasse um movimento na plataforma para que se testasse o funcionamento do algoritmo.

Pouco tempo após o incidente, Liz Kelley, funcionária da empresa se pronunciou no site:

Obrigado a todos que levantaram isso. testamos o preconceito antes de enviar o modelo e não encontramos evidências de preconceito racial ou de gênero em nossos testes, mas está claro que temos mais análises a fazer. Abriremos o código do nosso trabalho para que outros possam revisar e replicar.²⁸

²⁷ Tradução livre. MIDLAND, C. A faculty member has been asking how to stop Zoom from removing his head when he uses a virtual background. We suggested the usual plain background, good lighting etc, but it didn't work. I was in a meeting with him today when I realized why it was happening. 2020. Twitter. Disponível em: <https://twitter.com/colinmadland/status/1307111816250748933?s=20&t=a4iADcNCjtpdsFuPrOIMYg>. Acesso em: 01/06/2022.

²⁸ Tradução livre. KELLEY, L. thanks to everyone who raised this. we tested for bias before shipping the model and didn't find evidence of racial or gender bias in our testing, but it's clear that we've got more analysis to do. we'll open source our work so others can review and replicate. 2020. Twitter. Disponível em: <https://twitter.com/lizkelley/status/1307742267193532416>. Acesso em: 02/06/2022.

Investigações a respeito do algoritmo foram feitas por um período de um ano pela empresa, e, em setembro de 2021, os estudos²⁹ a respeito do caso foram publicados e se constatou a mudança na abordagem do algoritmo.

Outro caso relacionado à empresa surgiu em maio deste ano, onde Elon Musk, em um dos seus *tweets*, menciona:

Muito importante corrigir o seu feed do Twitter:

1. Toque no botão *home*
2. Toque nas estrelas no canto superior direito da tela
3. Selecione “Últimos *tweets*”

Você está sendo manipulado pelo algoritmo de maneiras que não percebe
É fácil de alternar entre as opções para ver a diferença.³⁰

Musk aborda o fato de que algoritmos podem manipular informações que são mostradas no *feed* de notícias dos usuários sem que eles percebam, e em seguida complementa:

Não estou sugerindo malícia no algoritmo, mas sim que ele está tentando adivinhar o que você pode querer ler e, ao fazê-lo, manipular/amplificar inadvertidamente seus pontos de vista sem que você perceba que isso está acontecendo³¹

Por intermédio de seus *tweets*, Musk levanta a indagação de que algoritmos possam estar sendo usados para manipulação de informação, mesmo que indiretamente. Ao filtrar a informação destinada a um usuário e escolher o que mostrar, o algoritmo estaria manipulando os pontos de vista sobre determinados assuntos sem que o indivíduo consiga ver o quadro da situação, o que é referido atualmente como “filtro bolha”. Um filtro bolha, ou *filter bubble*, é um estado de isolamento intelectual resultado de pesquisas personalizadas quando um algoritmo valora seletivamente quais informações um usuário recebe com base nas informações sobre o usuário, como localização, comportamento de navegação e histórico de pesquisa. Como resultado, o usuário se separa das informações que discorda de seus pontos de vista, isolando-o efetivamente em sua própria bolha cultural ou ideológica. Logo após, Musk finaliza: “Sem mencionar possíveis bugs no código. O *open source* é o caminho a percorrer para resolver tanto a confiança quanto a eficácia.”³²

Em suma, podemos perceber que as ocorrências de discriminação possuem diferentes resultados e meios com que se realiza o processamento de dados, portanto devem ser tratados de forma diferente. Devemos ter cuidado ao programar algoritmos para que se observe

²⁹ YEA, K.; TANTIPONGPIPAT, U.; MISHRA, S. Image Cropping on Twitter: Fairness Metrics, their Limitations, and the Importance of Representation, Design, and Agency. 2021. Disponível em: <https://arxiv.org/pdf/2105.08667.pdf>.

³⁰ Tradução livre. MUSK, E. Very important to fix your Twitter feed: 1. Tap home button. 2. Tap stars on upper right of screen. 3. Select “Latest tweets”. You are being manipulated by the algorithm in ways you don’t realize. Easy to switch back & forth to see the difference. 2022. Twitter. Disponível em: <https://twitter.com/elonmusk/status/1525612988115320838?s=20&t=UJPhFzvvtDXpWKcDtc9giw>. Acesso em: 07/06/2022.

³¹ Tradução livre. MUSK, E. I’m not suggesting malice in the algorithm, but rather that it’s trying to guess what you might want to read and, in doing so, inadvertently manipulate/amplify your viewpoints without you realizing this is happening. 2022. Twitter. Disponível em: <https://twitter.com/elonmusk/status/1525738556102164480?s=20&t=LgwyzzLwcmDfaYrLY63uvw>. Acesso em: 08/06/2022.

³² Tradução livre. MUSK, E. Not to mention potential bugs in the code. Open source is the way to go to solve both trust and efficacy. 2022. Twitter. Disponível em: <https://twitter.com/elonmusk/status/1525739780323016704?s=20&t=66WNLp13U7Mw0TRNJ86qmg>. Acesso em: 08/06/2022.

princípios básicos de funcionamento da sociedade, seja garantido ao indivíduo o tratamento justo com base nas garantias fundamentais e, ao mesmo tempo, tenha transparência no seu código.

4 REFLEXO NORMATIVO

Eventos, ou acontecimentos que marcam a história ou modificam a realidade, em sua grande maioria, são refletidos no ordenamento jurídico, que, afinal, nada mais é que o reflexo da sociedade. No caso da discriminação algorítmica, temos dois caminhos lógicos a seguir, o primeiro na garantia fundamental do sistema jurídico à não discriminação, e o segundo, no reflexo advindo da problemática da evolução tecnológica, que ficou positivado na Lei Geral de Proteção de Dados Pessoais (LGPD).

4.1 DISCRIMINAÇÃO NO ORDENAMENTO JURÍDICO

A proteção positiva contra discriminação é muito presente no ordenamento jurídico brasileiro, ela é versada em disposições constitucionais, tratados internacionais ratificados, assim como em legislações infraconstitucionais.

A Constituição da República, em seu artigo 3º, inciso IV afirma que “promover o bem de todos, sem preconceitos de origem, raça, sexo, cor, idade e quaisquer outras formas de discriminação” constitui um objetivo fundamental da República Federativa do Brasil.

No artigo 5º, em seu *caput*, aborda o princípio da igualdade, o qual garante que todos são iguais perante a lei e, portanto, não deve ocorrer discriminação de qualquer tipo. Assim, nos incisos XLI e XLII, estabelece a punição de qualquer tipo de discriminação que atente aos direitos e liberdades fundamentais e define o racismo como crime inafiançável e imprescritível, sujeito à pena de reclusão.

Em seu artigo 7º, inciso XXXI, em texto constitucional, fica expressa a proibição de qualquer discriminação no tocante a salário e critérios de admissão do trabalhador portador de deficiência.

Da mesma forma, em seu artigo 227:

É dever da família, da sociedade e do Estado assegurar à criança, ao adolescente e ao jovem, com absoluta prioridade, o direito à vida, à saúde, à alimentação, à educação, ao lazer, à profissionalização, à cultura, à dignidade, ao respeito, à liberdade e à convivência familiar e comunitária, além de colocá-los a salvo de toda forma de negligência, discriminação, exploração, violência, crueldade e opressão.³³

Exalta o dever da família, do estado e da sociedade de proteger a criança e o adolescente de toda forma de discriminação. Ademais, no parágrafo 1º, inciso II, estabelece o dever do Estado na criação de programas para extinção de todas as formas de discriminação para as pessoas portadoras de deficiência física. Em seu parágrafo 6º, proíbe a discriminação em relação aos filhos, havidos ou não da relação do casamento, ou por adoção.

Também se há a proteção jurídica infraconstitucional contra discriminação, como por exemplo, em tratados internacionais, incorporados à jurisdição por meio de decretos, como, por exemplo, o Decreto nº 65.810/ 1969, Convenção Internacional sobre a Eliminação de todas as Formas de Discriminação Racial, onde atribui o conceito de discriminação racial:

³³ BRASIL. CONSTITUIÇÃO DA REPÚBLICA FEDERATIVA DO BRASIL DE 1988, BRASÍLIA, 2022. Disponível em: http://www.pla.br/ccivil_03/constituicao/constituicao.html Acesso em: 15/06/2022.

Nesta Convenção, a expressão “discriminação racial” significará qualquer distinção, exclusão, restrição ou preferência baseadas em raça, cor, descendência ou origem nacional ou étnica que tem por objetivo ou efeito anular ou restringir o reconhecimento, gozo ou exercício num mesmo plano, (em igualdade de condição), de direitos humanos e liberdades fundamentais no domínio político econômico, social, cultural ou em qualquer outro domínio de vida pública.³⁴

Assim como o Decreto nº 3.956/2001, que promulgou a Convenção Interamericana para a Eliminação de Todas as Formas de Discriminação contra as Pessoas Portadoras de Deficiência, onde define o conceito de discriminação contra as pessoas portadoras de deficiência:

O termo "discriminação contra as pessoas portadoras de deficiência" significa toda diferenciação, exclusão ou restrição baseada em deficiência, antecedente de deficiência, conseqüência de deficiência anterior ou percepção de deficiência presente ou passada, que tenha o efeito ou propósito de impedir ou anular o reconhecimento, gozo ou exercício por parte das pessoas portadoras de deficiência de seus direitos humanos e suas liberdades fundamentais.³⁵

Temos também o Decreto nº 4.377/2002, que promulgou a Convenção sobre a Eliminação de Todas as Formas de Discriminação contra a Mulher, de 1979, e revogou o Decreto no 89.460, de 20 de março de 1984, que define o conceito de discriminação contra mulher:

Para os fins da presente Convenção, a expressão "discriminação contra a mulher" significará toda a distinção, exclusão ou restrição baseada no sexo e que tenha por objeto ou resultado prejudicar ou anular o reconhecimento, gozo ou exercício pela mulher, independentemente de seu estado civil, com base na igualdade do homem e da mulher, dos direitos humanos e liberdades fundamentais nos campos político, econômico, social, cultural e civil ou em qualquer outro campo.³⁶

Temos também, o Decreto nº 9.883/2019, que dispõe sobre o Conselho Nacional de Combate à Discriminação.

Por fim, o Decreto nº 10.932/2022, que promulgou a Convenção Interamericana contra o Racismo, a Discriminação Racial e Formas Correlatas de Intolerância, firmado pela República Federativa do Brasil, na Guatemala, em 5 de junho de 2013, onde define diversos conceitos de discriminação, como discriminação racial:

Discriminação racial é qualquer distinção, exclusão, restrição ou preferência, em qualquer área da vida pública ou privada, cujo propósito ou efeito seja anular ou restringir o reconhecimento, gozo ou exercício, em condições de igualdade, de um ou mais direitos humanos e liberdades fundamentais consagrados nos instrumentos

³⁴ BRASIL. DECRETO Nº 65.810, DE 8 DE DEZEMBRO DE 1969. Promulga a Convenção Internacional sobre a Eliminação de todas as Formas de Discriminação Racial., BRASÍLIA, 2022. Disponível em: http://www.planalto.gov.br/ccivil_03/decreto/1950-1969/D65810.html. Acesso em: 16/06/2022.

³⁵ BRASIL. DECRETO Nº 3.956, DE 8 DE OUTUBRO DE 2001. Promulga a Convenção Interamericana para a Eliminação de Todas as Formas de Discriminação contra as Pessoas Portadoras de Deficiência., BRASÍLIA, 2022. Disponível em: http://www.planalto.gov.br/ccivil_03/decreto/2001/D3956.htm. Acesso em: 17/06/2022.

³⁶ BRASIL. DECRETO Nº 4.377, DE 13 DE SETEMBRO DE 2002. Promulga a Convenção sobre a Eliminação de Todas as Formas de Discriminação contra a Mulher, de 1979, e revoga o Decreto no 89.460, de 20 de março de 1984., BRASÍLIA, 2022. Disponível em: http://www.planalto.gov.br/ccivil_03/decreto/2002/d4377.htm. Acesso em: 17/06/2022.

internacionais aplicáveis aos Estados Partes. A discriminação racial pode basear-se em raça, cor, ascendência ou origem nacional ou étnica.³⁷

Define também o conceito de discriminação racial indireta:

Discriminação racial indireta é aquela que ocorre, em qualquer esfera da vida pública ou privada, quando um dispositivo, prática ou critério aparentemente neutro tem a capacidade de acarretar uma desvantagem particular para pessoas pertencentes a um grupo específico, com base nas razões estabelecidas no Artigo 1.1, ou as coloca em desvantagem, a menos que esse dispositivo, prática ou critério tenha um objetivo ou justificativa razoável e legítima à luz do Direito Internacional dos Direitos Humanos.³⁸

Firma o conceito de discriminação múltipla:

Discriminação múltipla ou agravada é qualquer preferência, distinção, exclusão ou restrição baseada, de modo concomitante, em dois ou mais critérios dispostos no Artigo 1.1, ou outros reconhecidos em instrumentos internacionais, cujo objetivo ou resultado seja anular ou restringir o reconhecimento, gozo ou exercício, em condições de igualdade, de um ou mais direitos humanos e liberdades fundamentais consagrados nos instrumentos internacionais aplicáveis aos Estados Partes, em qualquer área da vida pública ou privada.³⁹

E aborda conceitos como racismo, intolerância e medidas especiais. Em relação a medidas especiais, que abordamos como discriminação positiva, afirma:

As medidas especiais ou de ação afirmativa adotadas com a finalidade de assegurar o gozo ou exercício, em condições de igualdade, de um ou mais direitos humanos e liberdades fundamentais de grupos que requeiram essa proteção não constituirão discriminação racial, desde que essas medidas não levem à manutenção de direitos separados para grupos diferentes e não se perpetuem uma vez alcançados seus objetivos.⁴⁰

Ademais, existem outros exemplos de normas jurídicas infraconstitucionais, porém possuem caráter punitivo ou preventivo de condutas discriminatórias, assegurando o direito de

³⁷ BRASIL. DECRETO Nº 10.932, DE 10 DE JANEIRO DE 2022. Promulga a Convenção Interamericana contra o Racismo, a Discriminação Racial e Formas Correlatas de Intolerância, firmado pela República Federativa do Brasil, na Guatemala, em 5 de junho de 2013, BRASÍLIA, 2022. Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2019-2022/2022/Decreto/D10932.htm. Acesso em: 17/06/2022.

³⁸ BRASIL. DECRETO Nº 10.932, DE 10 DE JANEIRO DE 2022. Promulga a Convenção Interamericana contra o Racismo, a Discriminação Racial e Formas Correlatas de Intolerância, firmado pela República Federativa do Brasil, na Guatemala, em 5 de junho de 2013, BRASÍLIA, 2022. Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2019-2022/2022/Decreto/D10932.htm. Acesso em: 17/06/2022.

³⁹ BRASIL. DECRETO Nº 10.932, DE 10 DE JANEIRO DE 2022. Promulga a Convenção Interamericana contra o Racismo, a Discriminação Racial e Formas Correlatas de Intolerância, firmado pela República Federativa do Brasil, na Guatemala, em 5 de junho de 2013, BRASÍLIA, 2022. Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2019-2022/2022/Decreto/D10932.htm. Acesso em: 17/06/2022.

⁴⁰ BRASIL. DECRETO Nº 10.932, DE 10 DE JANEIRO DE 2022. Promulga a Convenção Interamericana contra o Racismo, a Discriminação Racial e Formas Correlatas de Intolerância, firmado pela República Federativa do Brasil, na Guatemala, em 5 de junho de 2013, BRASÍLIA, 2022. Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2019-2022/2022/Decreto/D10932.htm. Acesso em: 17/06/2022.

igualdade. Como é o exemplo da Lei nº 7.716/1989 (que define os crimes resultantes de preconceito de raça ou de cor), da Lei nº 8.069/1990 (que dispõe sobre o Estatuto da Criança e do Adolescente e dá outras providências), da Lei nº 10.741/2003 (que dispõe sobre o Estatuto do Idoso e dá outras providências), da Lei nº 11.340/2006 (que cria mecanismos para coibir a violência doméstica e familiar contra a mulher), entre outras.

Embora o ordenamento jurídico garanta o direito à não discriminação, com o advento da inteligência artificial e algoritmos de aprendizado de máquina surgem outros problemas que precisam ser abordados. Para resolver essas adversidades, o Ordenamento Jurídico Brasileiro introduziu a Lei Geral de Proteção de Dados Pessoais.

4.2 A LEI GERAL DE PROTEÇÃO DE DADOS PESSOAIS (LGPD)

Com o surgimento dos novos problemas relacionados ao avanço tecnológico, a garantia fundamental de não discriminação não foi suficiente para assegurar direitos fundamentais quanto ao uso de tecnologia, sendo necessária a introdução da Lei Geral de Proteção de Dados que abordou tanto o direito à privacidade como também direitos e liberdades fundamentais.

O tráfego de dados pessoais por empresas informativas cresce cada vez mais. Essas informações são coletadas e vendidas sem sequer o titular dos dados saber, ou ter controle da utilização. Nesse sentido, a Lei nº 13.709/2018 (Lei Geral de Proteção de Dados Pessoais), espelhada na legislação europeia a respeito do tema, a *General Data Protection Regulation* (GDPR), foi implementada com objetivo de garantir a proteção de dados e informações dos indivíduos da sociedade.

A Lei Geral de Proteção de Dados é fácil de se compreender, pois é uma legislação conceitual e principiológica. Ou seja, a Lei elenca e define conceitos que se fazem necessários à compreensão para abordagem da lei, como também estabelece e enumera princípios fundamentais que fazem parte da estrutura da lei. A lei traz como seus fundamentos, no art. 2º:

- I - o respeito à privacidade;
- II - a autodeterminação informativa;
- III - a liberdade de expressão, de informação, de comunicação e de opinião;
- IV - a inviolabilidade da intimidade, da honra e da imagem;
- V - o desenvolvimento econômico e tecnológico e a inovação;
- VI - a livre iniciativa, a livre concorrência e a defesa do consumidor; e
- VII - os direitos humanos, o livre desenvolvimento da personalidade, a dignidade e o exercício da cidadania pelas pessoas naturais.⁴¹

Nota-se, que a proteção dos direitos fundamentais é bastante evidente no artigo 2º da LGPD, que pode ser relacionado ao texto constitucional brasileiro no que concerne ao conteúdo, haja vista que a Constituição Federal Brasileira é pautada nos direitos fundamentais (PINHEIRO, 2020).

Logo após, no artigo 5º, a lei aborda e conceitua os termos que serão utilizados, dentre eles, temos o conceito de dados pessoais sensíveis que define como:

dado pessoal sobre origem racial ou étnica, convicção religiosa, opinião política, filiação a sindicato ou a organização de caráter religioso, filosófico ou

⁴¹ BRASIL. LEI Nº 13.709, DE 14 DE AGOSTO DE 2018. Lei Geral de Proteção de Dados Pessoais (LGPD), BRASÍLIA, 2022. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/113709.htm. Acesso em: 17/06/2022.

político, dado referente à saúde ou à vida sexual, dado genético ou biométrico, quando vinculado a uma pessoa natural.⁴²

Tal conceito se faz relevante na abordagem do tema do presente trabalho, pois exemplificamos casos em que dados pessoais sensíveis eram utilizados como critérios em sistemas de algoritmos, sendo uma prática discriminatória.

Em seguida, a Lei define o que se entende por consentimento: “manifestação livre, informada e inequívoca pela qual o titular concorda com o tratamento de seus dados pessoais para uma finalidade determinada” (BRASIL, 2018). Definição também relevante ao presente artigo, pois em alguns algoritmos de inteligência artificial, dados pessoais são usados sem o fornecimento de consentimento pelo titular, algo que a Lei abordada visa regular em seu artigo 7º, inciso I.

Com relação ao fato de ser uma legislação principiológica, a Lei aborda, em seu artigo 6º, os princípios que ordenam todo o conteúdo da presente lei:

I - finalidade: realização do tratamento para propósitos legítimos, específicos, explícitos e informados ao titular, sem possibilidade de tratamento posterior de forma incompatível com essas finalidades;

II - adequação: compatibilidade do tratamento com as finalidades informadas ao titular, de acordo com o contexto do tratamento;

III - necessidade: limitação do tratamento ao mínimo necessário para a realização de suas finalidades, com abrangência dos dados pertinentes, proporcionais e não excessivos em relação às finalidades do tratamento de dados;

IV - livre acesso: garantia, aos titulares, de consulta facilitada e gratuita sobre a forma e a duração do tratamento, bem como sobre a integralidade de seus dados pessoais;

V - qualidade dos dados: garantia, aos titulares, de exatidão, clareza, relevância e atualização dos dados, de acordo com a necessidade e para o cumprimento da finalidade de seu tratamento;

VI - transparência: garantia, aos titulares, de informações claras, precisas e facilmente acessíveis sobre a realização do tratamento e os respectivos agentes de tratamento, observados os segredos comercial e industrial;

VII - segurança: utilização de medidas técnicas e administrativas aptas a proteger os dados pessoais de acessos não autorizados e de situações acidentais ou ilícitas de destruição, perda, alteração, comunicação ou difusão;

VIII - prevenção: adoção de medidas para prevenir a ocorrência de danos em virtude do tratamento de dados pessoais;

IX - não discriminação: impossibilidade de realização do tratamento para fins discriminatórios ilícitos ou abusivos;

X - responsabilização e prestação de contas: demonstração, pelo agente, da adoção de medidas eficazes e capazes de comprovar a observância e o cumprimento das normas de proteção de dados pessoais e, inclusive, da eficácia dessas medidas.⁴³

Dentre eles, se fazem relevantes ao presente artigo os princípios: da “finalidade”, da “transparência” e o da “não discriminação”.

Quanto ao princípio da finalidade, o legislador exemplifica que a utilização dos dados deve ter finalidade específica, explícita e deve ser informado ao titular dos dados a coleta dos

⁴² BRASIL. LEI Nº 13.709, DE 14 DE AGOSTO DE 2018. Lei Geral de Proteção de Dados Pessoais (LGPD), BRASÍLIA, 2022. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/113709.htm. Acesso em: 17/06/2022.

⁴³ BRASIL. LEI Nº 13.709, DE 14 DE AGOSTO DE 2018. Lei Geral de Proteção de Dados Pessoais (LGPD), BRASÍLIA, 2022. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/113709.htm. Acesso em: 17/06/2022.

dados, assim como para que se destinam, não podendo direcionar os dados a outra finalidade senão àquela informada. Esse princípio dificulta a troca e venda de dados entre empresas, pois não há de se informar onde e como os dados vão ser utilizados.

Logo após, temos o princípio da transparência que garante, aos titulares de dados, informações claras, precisas e facilmente acessíveis. É um princípio fundamental para abordagem dos problemas advindos de algoritmos de inteligência artificial, pois se precisa ter acesso ao uso de seus dados pelos algoritmos, além de examinar como foram utilizados os seus dados. Afinal, não há como afirmar se a conduta foi lícita ou ilícita se não tiver como averiguar como foi dada a tomada de decisão. Em seu artigo 20, a Lei estabelece em seu parágrafo 1º e 2º, que o controlador deverá fornecer, sempre que solicitado, informações claras e adequadas a respeito dos critérios e dos procedimentos utilizados para a decisão automatizadas, e, se não disponibilizadas, a autoridade nacional poderá realizar auditoria para verificação de aspectos discriminatórios.

Por fim, o princípio da não discriminação que veda a realização do tratamento de dados pessoais para fins discriminatórios ilícitos ou abusivos. O inciso exemplifica o conceito trabalhado de discriminação negativa, porém não proíbe o que vimos como discriminação positiva.

Faz-se importante ressaltar que, em seu artigo 20, a Lei garante o direito a solicitar a revisão de decisões tomadas unicamente com base em tratamento automatizado de dados pessoais que afetem seus interesses. Portanto, o titular do dado que se sentir prejudicado por uma decisão automatizada, pode solicitar uma revisão da decisão e pedir explicação a respeito dos critérios e procedimentos utilizados para a tomada de decisão do sistema de inteligência artificial.

Portanto, pode se concluir que a Lei Geral de Proteção de Dados Pessoais estabelece diretrizes importantes para o tratamento de dados pessoais e oferece garantia ampla quanto aos princípios fundamentais de direito.

5 CONSIDERAÇÕES FINAIS

A velocidade com que a sociedade contemporânea produz e aperfeiçoa avanços tecnológicos é sem precedentes. A realidade caminha cada vez mais para um mundo governado pela tecnologia. Com o aumento descomunal de dados gerados, algoritmos de inteligência artificial vêm sendo desenvolvidos com extrema rapidez. Com o intuito de facilitar o processamento de dados, que crescem exponencialmente, algoritmos de inteligência artificial automatizam decisões e governam aspectos cotidianos de indivíduos, apontando resultados em uma pesquisa, indicando quais filmes serão recomendados, selecionando notícias e fatos que serão exibidos, revelando quais fotos serão exibidas em sites de redes sociais até decisões de mais impacto pessoal, como a aprovação de um empréstimo bancário ou de uma vaga de emprego.

Dada a escala e complexidade da sociedade atual, bem como nosso conhecimento sobre a natureza humana, seria extremamente ingênuo confiar simplesmente no comportamento ético espontâneo e voluntário de indivíduos e empresas para garantir a justiça ou melhorar a dignidade humana. Em uma realidade onde o lucro guia todas as decisões de empresas, e é cada vez mais procurado por elas, faz-se de extrema importância o papel da fiscalização e prevenção dos sistemas, assim como a garantia da defesa dos direitos fundamentais e subjetivos ocasionalmente derrubados ou revogados pelos sistemas. Algoritmos de inteligência artificial, quando usados corretamente, providenciam uma assistência fundamental para a sociedade. A capacidade humana não é suficiente para lidar com tamanha quantidade de dados gerados

atualmente, algoritmos ajudam a resolver esse problema essencial. Porém, nem sempre são utilizados de maneira correta, gerando consequências catastróficas para a sociedade.

Quando usados de maneira errada, algoritmos de inteligência artificial geram diversos questionamentos e problemas que não eram visionados no momento de sua criação por, em teoria, não possuem a capacidade de discernimento humano, não seriam capazes de produzir decisões intencionalmente maléficas. Isso fez com que os algoritmos tivessem uma lógica presumida de boa-fé, dificultando o diálogo a respeito da discriminação algorítmica. Embora seja verdadeiro que algoritmos não possuem consciência e intenção de discriminar, quando abastecidos com dados viciados, geram resultados enviesados, e, muitas vezes, acabam reproduzindo o erro que foram programados para evitar. Além disso, em algumas situações, seu uso é fundamentalmente discriminatório. Quando os valores que o algoritmo usa para chegar a um resultado são baseados somente em critérios de identidade de grupo, o resultado é fundamentalmente discriminatório. Contudo, vale salientar que o simples fato de haver discrepâncias no resultado não evidencia automaticamente a discriminação do algoritmo, há de se provar de fato a intenção ou o enviesamento do programa. Ademais, outros problemas surgem da má implementação dos algoritmos, como o caso do reconhecimento facial trazido anteriormente. Pelo abastecimento de dados predominantemente de uma raça, o algoritmo falhava e não detectava o rosto de outras raças, levando à exclusão de participação de acesso para as outras etnias. Temos também o caso de manipulação de informações, que acaba criando zonas de isolamento intelectual e cultural em indivíduos que deixam de receber informações isentas.

Evidentemente, o ordenamento jurídico já possuía ferramentas para lidar com a discriminação advinda do tratamento desigual por algoritmos de inteligência artificial, como enumerados diversos dispositivos constitucionais e infraconstitucionais. Porém, não se tinha a segurança jurídica de prevenção e fiscalização desses sistemas. Com a promulgação da LGPD, o indivíduo ficou mais resguardado em relação a esses problemas.

Entretanto, estamos caminhando rapidamente para um cenário hipotético de superinteligência artificial que excede em muito a capacidade intelectual de qualquer ser humano. Se concluirmos que a inteligência artificial pode vir a influenciar de forma negativa o público, se faz evidente o fato de que temos que ter regulamentos na sua programação e fiscalização em sua aplicação.

Dito isso, é inegável que a inteligência artificial já tenha trazido grandes benefícios para a sociedade e, negar isso, não é o objetivo do trabalho, justamente o contrário. A inteligência artificial tem o potencial de melhorar completamente o padrão de vida da sociedade e tornar o mundo mais justo e igualitário.

Portanto, o que se pôde inferir do artigo é que o rápido avanço das tecnologias de inteligência artificial, como os algoritmos de aprendizado de máquina, têm causado alterações relevantes no quadro político, social e econômico da sociedade. E instrumentos como a LGPD são um passo na direção certa para a regulamentação e fiscalização dos sistemas. Ademais, o intuito principal do presente escrito é evidenciar a necessidade de educar e capacitar o operador de Direito para enfrentar os sistemas que infringem direitos fundamentais e subjetivos em função do enviesamento do próprio algoritmo, com atuação proativa junto à sociedade e às instituições, para que se possa garantir os direitos fundamentais e não frear o desenvolvimento tecnológico.

REFERÊNCIAS

- ANYOHA, Rockwell. BLOG, SPECIAL EDITION ON ARTIFICIAL INTELLIGENCE. **The History of Artificial Intelligence**, [s. l.], 28 ago. 2017.
- BRASIL. CONSTITUIÇÃO DA REPÚBLICA FEDERATIVA DO BRASIL DE 1988, BRASÍLIA, 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 15/06/2022.
- BRASIL. DECRETO No 10.932, DE 10 DE JANEIRO DE 2022. **Promulga a Convenção Interamericana contra o Racismo, a Discriminação Racial e Formas Correlatas de Intolerância, firmado pela República Federativa do Brasil, na Guatemala, em 5 de junho de 2013**, BRASÍLIA, 2022. Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2019-2022/2022/Decreto/D10932.htm. Acesso em: 17/06/2022.
- BRASIL. DECRETO No 3.956, DE 8 DE OUTUBRO DE 2001. **Promulga a Convenção Interamericana para a Eliminação de Todas as Formas de Discriminação contra as Pessoas Portadoras de Deficiência**, BRASÍLIA, 2001. Disponível em: http://www.planalto.gov.br/ccivil_03/decreto/2001/D3956.htm. Acesso em: 17/06/2022.
- BRASIL. DECRETO No 4.377, DE 13 DE SETEMBRO DE 2002. **Promulga a Convenção sobre a Eliminação de Todas as Formas de Discriminação contra a Mulher, de 1979, e revoga o Decreto no 89.460, de 20 de março de 1984**, BRASÍLIA, 2002. Disponível em: http://www.planalto.gov.br/ccivil_03/decreto/2002/d4377.htm. Acesso em: 17/06/2022.
- BRASIL. DECRETO No 65.810, DE 8 DE DEZEMBRO DE 1969. **Promulga a Convenção Internacional sobre a Eliminação de todas as Formas de Discriminação Racial**, BRASÍLIA, 1969. Disponível em: http://www.planalto.gov.br/ccivil_03/decreto/1950-1969/D65810.html. Acesso em: 16/06/2022.
- BRASIL. LEI No 13.709, DE 14 DE AGOSTO DE 2018. **Lei Geral de Proteção de Dados Pessoais (LGPD)**, BRASÍLIA, 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/113709.htm. Acesso em: 17/06/2022.
- BROWN, R. **Prejudice: Its Social Psychology**. [S.l.]: Wiley-Blackwell, 2018.
- BUOLAMWINI, J. **Aspire Mirror**. Disponível em: <http://www.aspiremirror.com>. Acesso em: 01/06/2022.
- EQUALITY, C. for R. **Medical school admissions Report of a formal investigation into St George's Hospital Medical School**. Disponível em: <https://www.jstor.org/stable/pdf/community.28327674.pdf>. Acesso em: 15/05/2022.
- GOMES, J. B. **Ação Afirmativa e Princípio Constitucional da Igualdade**. Rio de Janeiro: Renovar, 2001
- HAENLEIN, Michael; KAPLAN, Andreas M. **A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence.**, [s. l.], 1 ago. 2019.

KELLEY, L. **thanks to everyone who raised this. we tested for bias before shipping the model and didn't find evidence of racial or gender bias in our testing, but it's clear that we've got more analysis to do. we'll open source our work so others can review and replicate.** 2020. Twitter. Disponível em: <https://twitter.com/lizkelley/status/1307742267193532416>. Acesso em: 02/06/2022.

KORDZAECH, N.; GHASEMAGHAEI, M. **Algorithmic bias: review, synthesis, and future research directions.** Disponível em: shorturl.at/sFIM1. Acesso em: 10/05/2022.

MCCARTHY, J. **“the science and engineering of making intelligent machines.”.** 1955.

MIDLAND, C. **A faculty member has been asking how to stop Zoom from removing his head when he uses a virtual background. We suggested the usual plain background, good lighting etc, but it didn't work. I was in a meeting with him today when I realized why it was happening.** Twitter. Disponível em: <https://twitter.com/colinmadland/status/1307111816250748933?s=20&t=a4iADcNCjtpdsFuPrOIMYg>. Acesso em: 01/06/2022.

MOREIRA, A. J. **O que é discriminação?** [S.l.]: Casa do Direito, 2017.

MUSK, E. **I'm not suggesting malice in the algorithm, but rather that it's trying to guess what you might want to read and, in doing so, inadvertently manipulate/amplify your viewpoints without you realizing this is happening.** 2022. Twitter. Disponível em: <https://twitter.com/elonmusk/status/152573856102164480?s=20&t=LgwyzzLwcmDfaYrLY63uvw>. Acesso em: 08/06/2022.

MUSK, E. **Not to mention potential bugs in the code. Open source is the way to go to solve both trust and efficacy.** 2022. Twitter. Disponível em: <https://twitter.com/elonmusk/status/1525739780323016704?s=20&t=66WNLp13U7Mw0TRNJ86qmg>. Acesso em: 08/06/2022.

MUSK, E. **Very important to fix your Twitter feed: 1. Tap home button. 2. Tap stars on upper right of screen. 3. Select “Latest tweets”. You are being manipulated by the algorithm in ways you don't realize. Easy to switch back & forth to see the difference.** Twitter. Disponível em: <https://twitter.com/elonmusk/status/1525612988115320838?s=20&t=UJPhFzvwtdXpWKcDtc9gjw>. Acesso em: 07/06/2022.

O'NEIL, C. **Algoritmos de Destruição em Massa.** [S.l.]: Editora Rua do Sabão, 2021.

PETERSON, J. **On the Psychological and Social Significance of Identity.** Disponível em: <https://www.jordanbpeterson.com/blog-posts/on-the-psychological-and-social-significance-of-identity/>. Acesso em: 12/05/2022.

RIBEIRO, V. C. **Todos os tipos de discriminação.** 2018. Disponível em: <https://valeriag.jusbrasil.com.br/artigos/583475306/todos-os-tipos-de-discriminacao>. Acesso em: 03/05/2022.

SAMUEL, A. L. **Some Studies in Machine Learning Using the Game of Checkers.** IBM

Journal of Research and Development, v. 3, n. 3, p. 210 – 229, 1959. Disponível em: <http://dx.doi.org/10.1147/rd.33.0210>. Acesso em: 18/04/2022.

SERVICES, N. Y. S. D. of F. **Report on Apple Card Investigation**. 2021. Disponível em: https://www.dfs.ny.gov/system/files/documents/2021/03/rpt_202103_apple_card_investigation.pdf. Acesso em: 25/05/2022.

SHAPIRO, B. WATCH: **Ben Shapiro Breaks Down The ‘Argument From Authority’**. Daily Wire, 2017. Disponível em: <https://www.dailywire.com/news/watch-ben-shapiro-breaks-down-argument-authority-hank-berrien>. Acesso em: 28/04/2022.

SOWELL, T. **Discriminação e disparidades**. [S.l.]: Record, 2019.

TURING, A. M. **Computing machinery and intelligence**. Mind, LIX, n. 236, p. 433 – 460, outubro 1950. Disponível em: <https://bit.ly/2PCToam>. Acesso em: 17/04/2022.

WATCHER, S. **Making algorithms accountable and explainable: the need for a legal framework**. Disponível em: <https://www.research.ox.ac.uk/article/2018-10-15-making-algorithms-accountable-and-explainable-the-need-for-a-legal-framework>.

WEIZENBAUM, J. **Computer Power and Human Reason: From Judgment to Calculation**. [S.l.]: W. H. Freeman and Company, 1976.

WOZNIAK, S. **The same thing happened to us. I got 10x the credit limit. We have no separate bank or credit card accounts or any separate assets. Hard to get to a human for a correction though. It’s big tech in 2019**. 2019. Twitter. Disponível em: https://twitter.com/stevewoz/status/1193330241478901760?ref_src=twsrc%5Etfw%7Ctwcamp%5Etweetembed%7Ctwterm%5E1193330241478901760%7Ctwgr%5E%7Ctwcon%5Es1_&ref_url=https%3A%2F%2Fwww.idropnews.com%2Fnews%2Fno-the-apple-card-does-not-discriminate-against-women%2F154988%2F. Acesso em: 20/05/2022.

YEA, K.; TANTIPONGPIPAT, U.; MISHRA, S. **Image Cropping on Twitter: Fairness Metrics, their Limitations, and the Importance of Representation, Design, and Agency**. 2021. Disponível em: <https://arxiv.org/pdf/2105.08667.pdf>. Acesso em: 05/06/2022.