

## Inteligências artificiais: existe ética na implementação de algoritmos?

*Artificial intelligence: Are there ethics in algorithm implementation?*

Celito de Bona<sup>1</sup> , Eduarda Arruda Schons<sup>2</sup> , Luiza Lopes-Flois<sup>3</sup> 

<sup>1</sup> Doutor em Direito pela UNISINOS; Mestre em Direito Negocial pela Universidade Estadual de Londrina; Especialista lato sensu em Direito Civil e Processo Civil pela Universidade Paranaense; Especialista lato sensu em Filosofia do Direito pela Universidade Estadual do Oeste do Paraná; Graduado em Direito pela Universidade Paranaense; Professor Efetivo do Curso de Direito - Universidade Estadual do Oeste do Paraná (UNIOESTE), Campus de Marechal Cândido Rondon. E-mail: celito.bona@unioeste.br.

<sup>2</sup> Acadêmica do Curso de Direito - Universidade Estadual do Oeste do Paraná (UNIOESTE), Campus de Marechal Cândido Rondon. E-mail: eduarda.schons@unioeste.br.

<sup>3</sup> Acadêmica do Curso de Direito - Universidade Estadual do Oeste do Paraná (UNIOESTE), Campus de Marechal Cândido Rondon. E-mail: luiza.flois@unioeste.br.

### RESUMO

Este trabalho visa analisar a implementação de algoritmos em inteligência artificial e a forma pela qual seus vieses podem perpetuar discursos discriminatórios como o racismo. O objetivo é demonstrar o racismo algorítmico e suas possíveis consequências para a sociedade por meio de uma abordagem qualitativa, utilizando procedimentos exploratórios e bibliográficos. Nesse contexto, busca-se evidenciar que o desenvolvimento da inteligência artificial precisa ser regulamentado a fim de proporcionar aos usuários maior transparência e representatividade nesse ambiente tecnológico. Assim, almeja-se alcançar a justiça algorítmica e suprimir qualquer forma de discriminação algorítmica.

**Palavras-chave:** Inteligência artificial. Algoritmos enviesados. Racismo algorítmico. Justiça algorítmica.

### ABSTRACT

This paper aims to analyze the implementation of algorithms in artificial intelligence and the way in which their biases can perpetuate discriminatory discourses such as racism. The goal is to demonstrate algorithmic racism and its possible consequences for society through a qualitative approach, using exploratory and bibliographic procedures. In this context, we seek to show that the development of artificial intelligence needs to be regulated in order to provide users with greater transparency and representativeness in this technological environment. Thus, the goal is to achieve algorithmic justice and suppress any form of algorithmic discrimination.

**Keywords:** Artificial intelligence. Biased algorithms. Algorithmic racism. Algorithmic justice.

## 1 Introdução

A presente pesquisa visou a análise da ética<sup>1</sup> na implementação de algoritmos nas inteligências artificiais, na relação estabelecida entre o desenvolvimento de dados e a discriminação algorítmica por questões raciais. Partiu-se do questionamento da perpetuação de discursos preconceituosos no meio digital a partir da configuração de algoritmos e as implicações éticas de sua aplicação e os meios para mitigar eventuais softwares enviesados.

Nesse sentido, o objeto da pesquisa foi a apuração da relação entre a ética e os algoritmos das inteligências artificiais, tendo como objetivos específicos a conceituação de algoritmos, inteligência artificial, vieses, justiça algorítmica, ética algorítmica e racismo algorítmico. Além disso, demais propósitos foram a verificação do preconceito racial na construção das inteligências artificiais, seus efeitos negativos e os meios pelos quais se pode evitar sua ocorrência.

A pesquisa está organizada em quatro tópicos principais: no primeiro tópico, foram analisados os conceitos de algoritmos e inteligência artificial e a forma pela qual ocorre sua implementação. Na segunda parte, evidenciou-se o fenômeno da discriminação algorítmica, que acontece por meio de vieses implantados nos algoritmos, e exemplificou-se esse cenário. No terceiro tópico, expôs-se as consequências de tal racismo algorítmico, em relação ao reconhecimento facial que presumia que pessoas negras eram mais perigosas, e a desconfiança que se passa a ter sobre a tecnologia. E, na quarta parte, foram introduzidas propostas para solucionar a problemática, com a justiça e ética algorítmicas e a atuação da legislação brasileira.

Assim, buscou-se demonstrar que as inteligências artificiais podem ser veículo de disseminação de preconceitos, como foi apresentado pelos diversos exemplos de racismo algorítmico. A problemática tem sido tratada como prioridade na atualidade pelas empresas

---

<sup>1</sup> Para os fins deste trabalho, considerando as várias concepções sobre ética existentes, será adotada a defendida por Ronald Dworkin sobre uma comunidade de princípios, definida em “O Império do Direito”, na qual se pressupõe que “[...] o Estado aja segundo um conjunto único e coerente de princípios mesmo quando seus cidadãos estão divididos quanto a natureza exata dos princípios de justiça e equidade corretos [...], admitimos a possibilidade de reconhecer que os atos das outras pessoas expressam uma concepção de equidade, justiça ou decência [...]. Essa capacidade é uma parte importante de nossa capacidade mais geral de tratar os outros com respeito, sendo, portanto, um requisito prévio de civilização” (DWORKIN, 1999, p. 202).

desenvolvedoras de softwares, pelos governos que buscam regular a prática e uso ético e pelos indivíduos, que se preocupam pela concretização de seus direitos fundamentais no meio digital.

## 2 Algoritmos e inteligência artificial

Algoritmos podem ser definidos como uma sequência de instruções determinada para solucionar um problema, por meio de cálculos matemáticos, e são a base do processo de desenvolvimento de um software quando aplicados à lógica computacional, como conclui Gillespie (2018). Um software, por sua vez, é um conjunto de instruções que define a atuação do computador, no qual a partir da inserção de dados (*input*), obtêm-se um resultado (*output*). Assim, algoritmos são ferramentas responsáveis pelo processamento de uma quantidade extensa de dados, de forma a fracionar processos para que sejam realizados em um computador. Abreu *et al* (2022) destacam-nos por serem dos mais diversos tipos, sempre se pautando por um objetivo específico, seja pela determinação da melhor rota a ser tomada pelo usuário, como no Google Maps; pela determinação de ranking de vídeos mais visualizados segundo o tempo médio de visualização, como no Youtube; ou ainda pelo oferecimento da melhor experiência ao usuário, no Instagram, elencando-se na temporalidade e engajamento das postagens e contas.

Os algoritmos são componentes cruciais para o entendimento das inteligências artificiais, pois os utilizam no processo de execução. Por mais que a Inteligência Artificial (IA) não tenha um conceito universal que a defina, ela pode ser entendida como a reprodução de comportamentos tipicamente humanos por máquinas, alicerçada na manipulação dos algoritmos. Ela é aplicada em três áreas atualmente: *machine learning* (ou aprendizado de máquina), *deep learning* e *natural language processing*, como demonstra Tacca e Rocha (2018). Os autores exemplificam que esta última área é, em sua realidade, um sistema no qual os computadores examinam e entendem a fala, traduzindo e analisando sentimentos. O *deep learning*, todavia, é um método mais sofisticado, que mostra a percepção, por parte da máquina, de comportamentos e padrões, para os quais o próprio sistema relaciona uma solução, assemelhando-se em muito com as funções cognitivas humanas. O *machine learning*, por sua

vez, é um procedimento pelo qual as máquinas aprendem a partir de algoritmos e dados pelos quais foram configuradas.

Sendo as máquinas capazes de adquirir conhecimento por meio de experiências das quais não foram programadas e de se adaptar ao meio ao qual estão inseridas, é o sistema mais aplicado na atualidade. A máquina aprende enquanto trabalha, programando a si própria, o que pode acarretar resultados inesperados e, por vezes, preconceituosos. Exemplo claro dessa problemática, conforme Abreu *et al* (2022), é a inteligência artificial *Tay*, desenvolvida pela Microsoft, que, em menos de 24h de contato com os usuários, se ‘tornou racista’, isto é, reproduziu os comportamentos preconceituosos deles, por ter aprendido que isso era aceitável.

Erros na programação dos algoritmos, que possibilitem a ocorrência de fenômenos racistas na execução das IAs devem ser analisados a fim de que se entenda como se originaram. Para que se tenha instaurado o aprendizado de máquina, fez-se necessário a inserção de dados nesse sistema, que podem ter sido enviesados ou restritos, representando apenas uma parte da população. Ademais, esses *inputs* foram inseridos por uma pessoa – um programador – e não por uma máquina, o que pode sugerir a possibilidade de ter ocorrido erros nesse processo envolvendo algum parâmetro criado pelo profissional que permitiu uma conclusão racista.

### **3 Discriminação algorítmica**

As Inteligências Artificiais (IAs) têm se popularizado cada vez mais pela internet, pois oferecem rapidez e facilidade para a realização das mais variadas tarefas. Vale pontuar que, para se criar um modelo de IA, os programadores filtram e selecionam quais serão as informações que serão fornecidas para o sistema para que este as use para prever soluções futuras. Dessa forma, destaca-se que, para uma inteligência artificial de fato funcionar, ela irá depender de representações abstratas de determinados conteúdos que estão sob responsabilidade do programador de inseri-las. A qualidade de dados (conteúdos) que são informados ao sistema é de extrema importância, pois serão responsáveis diretamente pelo resultado. Este resultado tende a refletir o contexto social e as opiniões e perspectivas dos criadores e que são permeadas por discriminações, exclusões, etc.

As decisões algorítmicas, atualmente, estão inseridas direta ou indiretamente em diversos momentos e situações do cotidiano das pessoas, desde sistemas que fazem avaliações para liberação ou não de empréstimos pessoais a decisões judiciais. O desenvolvimento tecnológico das IAs não tem sido acompanhado no campo jurídico em quesito de regulação de algoritmos destinados aos seus devidos fins, dessa forma, as decisões tomadas por esses sistemas têm se exposto a impactar negativamente sobre os indivíduos, principalmente nos grupos sociais minoritários.

A título de exemplo, cita-se o *Beauty.ai*, o primeiro concurso de beleza no mundo em que os candidatos eram avaliados por inteligências artificiais. Os juízes foram constituídos por cinco diferentes algoritmos, e sua base de dados foi alimentada com imagens de pessoas consideradas ‘as mais bonitas’ pelos desenvolvedores e que, em um concurso real, seriam as prováveis vencedoras. Conforme o jornal Olhar Digital (2016), dentre os 6 mil candidatos, de mais de 100 diferentes países, que voluntariamente enviaram fotos dos seus rostos para o concurso, as IAs escolheram 44 vencedores para o concurso na qual um era negro e os outros 43, brancos. Dessa forma, explicita-se que o sistema tende a sempre reproduzir os comportamentos humanos, nesse caso, a discriminação, já que foi programado tendo como base pessoas brancas como as mais bonitas. Alex Zhavoronkov (2016), chefe de pesquisa da empresa, ressalta que o maior número de participantes brancos também foi fator para que a IA selecionasse predominantemente estes como vencedores.

O exemplo citado se encaixa de forma adequada no que se denomina de discriminação algorítmica, termo que comprova a existência do racismo na programação das inteligências artificiais. Como parte do processo de criação das IAs, nota-se que é possível surgir distorções do sistema. Essas alterações sistemáticas são conhecidas como vieses algorítmicos que tendem a surgir em algoritmos, devido à qualidade de dados que foram usados, à escolha de variáveis usadas no modelo e às suposições feitas pelos desenvolvedores desse algoritmo. Além do mais, os algoritmos poderão e irão se adaptar aos vieses sociais implícitos e explícitos em que ele foi exposto e que, conseqüentemente, geram estereótipos para as pessoas.

Nesse contexto, os vieses podem ser classificados de acordo com a fonte de origem – dados de treino, foco algorítmico, processamento algorítmico, transferência de contexto e interpretação – e são entendidos da seguinte forma:

*Ciências Sociais Aplicadas em Revista*, v. 25, n. 45, p. 225-246, Edição especial, 2023.

- a. Vieses por conta dos dados de entrada e que ocorrem devido às imagens que foram utilizadas, uma vez que elas podem variar em quesitos de qualidade, iluminação e ponto de vista, por exemplo.
- b. Vieses por conta do foco algorítmico e que ocorrem quando alguns dados não são utilizados, pois possuem restrições legais e acabam ficando ‘presos’ a determinados dados.
- c. Vieses por conta do processamento algorítmico e que ocorrem quando o próprio algoritmo acaba sendo enviesado, seja devido a alguma limitação própria ou às escolhas feitas pelos desenvolvedores.
- d. Vieses por conta da transferência de contexto e que ocorrem quando o algoritmo é colocado fora do contexto de sua real finalidade.
- e. Vieses por conta de interpretação e que ocorrem quando os dados inseridos para moldar os algoritmos não são coincidentes de fato com a aplicação.

Nesse sentido, os vieses implícitos na inteligência artificial, principalmente no *machine learning*, têm suscitado a emergência de preconceitos já existentes na sociedade no meio digital. A discriminação ocorre quando os algoritmos são treinados com exemplos que lhe são apresentados, por humanos, participantes de uma sociedade racista. Assim, entende-se essa discriminação algorítmica muito mais como um racismo algorítmico, como afirma Silva (2022): “o modo pelo qual as tecnologias reproduzem imaginários sociais e técnicos, fortalecendo a ordenação racializada de conhecimentos, recursos, espaço e violência em detrimento de grupos não-brancos” (SILVA, 2022, p. 69), por atingir sempre uma minoria socialmente estigmatizada e historicamente oprimida.

Ocorrências racistas por meio da ação de algoritmos não podem ser banalizadas e perpetuam discriminações, como, infelizmente, se percebe atualmente. Em 2015, o *Google Photos* classificou, em uma de suas etiquetas de sugestão de conteúdo, a foto de um grupo de amigos negros como gorilas (COSTA *et al*, 2023, *on-line*). Pergunta-se, então, quais dados foram utilizados para programar esse sistema; se houve, de fato, a representação de pessoas de diferentes etnias; ou se analisou imagens de pessoas de pele clara, em sua maioria, o que impossibilitou a máquina de reconhecer um homem negro como tal.

Como retrato do racismo algorítmico, há o *COMPAS (Correctional Offender Management Profiling for Alternative Sanctions* - Perfil de Gerenciamento Corretivo de Infratores para Sanções Alternativas), um sistema de IA usado pelos tribunais nos Estados Unidos para determinar a probabilidade de reincidência de um réu, como relata Soares *et al* (2022). Os critérios analisados, como local de residência, envolvimento anterior com drogas, histórico familiar e escolar, viabilizaram que pessoas negras fossem classificadas com 'alto risco' de reincidência, com muito mais facilidade e em maior quantidade, do que as pessoas brancas. Esse cenário corrobora os vieses implantados nos algoritmos, que, a partir de parâmetros inseridos pelos programadores, propagam comportamentos preconceituosos. Logo, se as linhas de programação são elaboradas por mãos humanas, seus erros também devem ser consertados por este mesmo meio.

## **4 Impacto social dos vieses algorítmicos**

### **4.1 'Tendências criminosas'**

O reconhecimento facial é formado pela IA, e os algoritmos desenvolvidos para detectar padrões faciais vêm sendo aplicados nas políticas públicas como uma forma de garantir a segurança da sociedade. Porém, na prática, essa tecnologia tem sido um meio para o desenvolvimento de encarceramento arbitrário em massa, com a violação de direitos fundamentais.

Segundo a Rede de Observatório da Segurança (2022), mais de 90% das prisões foram de pessoas negras identificadas por reconhecimento facial, mesmo que algumas dessas nunca tenham tido qualquer tipo de passagem pela polícia. O problema é que os dados são orientados por algoritmos de *Machine Learning* treinados a partir dos dados de crimes, no qual buscam padrões em informações como envolvimento social, o lugar em que reside, antecedentes, escolaridade etc. Porém, sabe-se que não há como afirmar que determinada pessoa irá cometer o crime que outra pessoa cometeu com a justificativa de possuir o mesmo perfil.

No caso de pessoas negras, além do racismo já evidente no país, a acusação destes pelo reconhecimento facial baseado em um 'perfil criminoso' apenas reforça estereótipos e discriminações. O art. 1º, III da CF/88 dispõe que: "A República Federativa do Brasil, formada pela união indissolúvel dos Estados e Municípios e do Distrito Federal, constitui-se em Estado

*Ciências Sociais Aplicadas em Revista, v. 25, n. 45, p. 225-246, Edição especial, 2023.*

Democrático de Direito e tem como fundamentos: III - a dignidade da pessoa humana”. Além disso, o art. 3º, I e IV da CF/88 dispõe que:

Constituem objetivos fundamentais da República Federativa do Brasil: I - construir uma sociedade livre, justa e solidária; [...] IV - promover o bem de todos, sem preconceitos de origem, raça, sexo, cor, idade e quaisquer outras formas de discriminação (BRASIL, 1988, Art. 3).

Com isso, é evidente que a utilização de algoritmos com viés de caráter criminal em relação às pessoas negras viola os princípios fundamentais estabelecidos pela Constituição Federal do Brasil. A persistência desses vieses reforça os estereótipos e discriminações, perpetuando as desigualdades históricas e acentuando o racismo estrutural ainda presente no país. Diante disso, reflete-se que as minorias e as comunidades mais marginalizadas, historicamente, são os principais alvos que a lei atinge desigualmente, sendo os mais vulneráveis a receberem pontuações de reincidência. Isso tende a levar a sentimentos de injustiça, desamparo e marginalização, que afetam negativamente a qualidade de vida e o bem-estar, principalmente, desse grupo, além de perpetuar representações negativas ou estereotipadas.

Ainda sobre a temática, no ano de 2019, conforme o próprio Supremo Tribunal Federal (STF), o ministro Alexandre de Moraes concedeu o Habeas Corpus (HC nº 172606-SP) para que se anulasse uma condenação de crime de roubo, pois a decisão condenatória havia sido apoiada no reconhecimento do réu em somente uma foto. No processo em questão, o fato de o reconhecimento fotográfico não ter sido confirmado pelas testemunhas durante a instrução processual, além de ausência de outros elementos de prova concretos, levantou dúvidas quanto à veracidade das acusações.

Como apresentado, as IAs de reconhecimento facial têm sido utilizadas em diversas áreas e momentos da vida de toda a população, principalmente no que concerne à aplicação da lei, porém, podem ser falhas e insuficientes para julgar alguém. Com a decisão de Moraes, ressalta-se a necessidade de haver uma prova consistente, hábil o suficiente para afastar qualquer incerteza, nesse caso, atingindo diretamente a fragilidade do reconhecimento fotográfico realizado.

Logo, é necessário que se garanta que o uso das IAs no campo jurídico (principalmente no que trata das imagens dos indivíduos) seja estruturado juntamente com outras provas e evidências confiáveis. A decisão do ministro do STF explicita a importância de um sistema judicial que exija provas consistentes, para que se evite condenações injustas fundamentadas em evidências insuficientes ou que não sejam precisas, como o reconhecimento facial.

### 3.2 Erosão da confiança na tecnologia

A maioria dos algoritmos de IA possui uma complexidade de compreensão para as pessoas no geral, criando um ambiente de opacidade na qual não há uma visibilidade sobre como as decisões são tomadas. A falta de transparência dificulta a responsabilização pelos vieses nos algoritmos, levando a questionamentos sobre justiça e igualdade.

Os casos de discriminação e preconceito que são evidentemente causados por algoritmos faz com que a confiança na tecnologia seja enfraquecida. O caso da discriminação ocorrido pela *Beauty.ai*, por exemplo, destaca a falha nos algoritmos e arruína a confiança de que a tecnologia seja utilizada de forma responsável.

A erosão da confiança na tecnologia também é causada devido à falta de diversidade e inclusão na equipe responsável pelo desenvolvimento de algoritmos, já que a ausência de uma equipe diversificada acaba por ter perspectivas limitadas e, assim, gerar propensões injustas que então se refletem nos algoritmos.

No que tange à falta de representatividade nesse meio, tem-se que os programadores, em sua maioria, são homens brancos, cis e heterossexuais, que, inconscientemente, transparecem sua visão de mundo ao inserirem dados passíveis de se tornarem enviesados. A ausência de mulheres negras nesse ramo de trabalho já foi criticada, e instituições como a Liga da Justiça Algorítmica tentam resolvê-la. Segundo Oliveira (2022), essa Liga foi criada pela pesquisadora Joy Buolamini, do MIT (*Massachusetts Institute of Technology*), ao perceber que os sistemas de reconhecimento facial estavam classificando mulheres negras como homens, visando o combate a esses preconceitos raciais e uma inclusão algorítmica.

A própria concentração de grandes corporações em uma restrita área geográfica já acarreta a exclusão de outras, em relação aos padrões sociais e culturais. Os Estados Unidos da América são o foco das empresas desenvolvedoras de IAs, que mais impactam no cotidiano das

pessoas, como Amazon, Facebook, Apple e Google e, por estarem agrupadas em um mesmo espaço, tornam-se reflexo umas das outras no que diz respeito às visões de mundo que compartilham.

Tal fenômeno pode ser exemplificado na exposição de mamilos femininos: redes sociais, como o Instagram, banem esse tipo de postagem, entendendo ser repudiável ao considerá-lo um conteúdo de nudez, com viés pornográfico. Contudo, essa é a visão cultural dos norte-americanos (generalizando, a título de exemplo) e, não, a de povos indígenas, que têm esse hábito como cultural e natural em seu cotidiano, bem como mulheres europeias, nos países em que é permitido *topless* em praias. Assim, essas fotografias são consideradas inadequadas em redes sociais, visto que estas representam um contexto sociocultural específico, enquanto excluem outros entendimentos.

As consequências dos vieses que são introduzidos nos algoritmos de IA são significativas, levando à resistência da adoção da tecnologia ou, em alguns casos, ainda, à rejeição. A falta de confiança nas tecnologias resulta na perda de oportunidades para inovação e melhoria nos serviços que a IA pode oferecer, deixando de aumentar o desenvolvimento em empresas, regiões e países.

## **5 Justiça algorítmica e mitigação de vieses discriminatórios**

Para que se possa obter um bom aproveitamento de toda a capacidade que as Inteligências Artificiais podem oferecer, é essencial investir no desenvolvimento contínuo e na análise constante dessas tecnologias, com o objetivo de evitar atos discriminatórios que possam prejudicar a sociedade. Dessa forma, para um desenvolvimento de forma responsável, é necessário a identificação e a mitigação dos vieses nos algoritmos.

Segundo a IBM (2021), atualmente existem ferramentas as quais foram criadas para ajudar as empresas a desenvolverem uma IA íntegra, como a *AI Fairness 360 Toolkit*, *AI FactSheets* e *IBM Watson OpenScale*, sendo, portanto, meios que ajudam na mitigação dos vieses.

### 5.1. Justiça algorítmica

A fim de combater a discriminação algorítmica dos modelos de inteligência artificial, torna-se crucial a mitigação de vieses nesses sistemas. Por esse motivo, busca-se uma equidade algorítmica, usando de algoritmos que não provoquem a continuidade de preconceitos no meio eletrônico nem ataquem a dignidade da pessoa humana e suas garantias fundamentais como cidadão para resolver questões sociais.

Entretanto, há diversas concepções de justiça que devem ser consideradas na aplicação de uma justiça algorítmica, pois não existe uma definição universal do que seria o mais justo, assim como também não existe definição de ética, mas apenas, métodos variados que almejam alcançar tais ideias. A não utilização de dados sensíveis nas IAs pode ser entendida como a concepção mais justa, o que resolveria a problemática do COMPAS, todavia, não poderia ser aplicada no caso do *Beauty.ai*, no qual o uso de uma menor quantidade de dados possibilitou a discriminação racial. Não há uma resposta certa para o que seria a justiça nesse caso, o que suscita discussões éticas.

A garantia de um tratamento equitativo, por algoritmos, é uma questão ética, mas também deve ser considerada como um direito do indivíduo. Em 2021, o Congresso Nacional chileno aprovou a Lei nº 21.383, que altera a redação do texto constituinte do país, o que colocou a proteção contra a discriminação algorítmica como um neurodireito; exemplo este que pode ser aplicado no Brasil, na inserção de uma nova geração de direitos fundamentais.

A necessidade de regulamentação da aplicação da inteligência artificial e da programação de algoritmos se evidencia para que o meio digital observe os princípios e objetivos fundamentais da Constituição Federal e os direitos fundamentais expostos na Declaração Universal dos Direitos Humanos. Logo, a conclusão lógica seria a criação de uma IA que obedecesse a um conjunto de regras que inibisse preconceitos e oferecesse tratamento igualitário para todas as pessoas.

Ex-executivos da *OpenAi*, desenvolvedora do *Chat GPT – chatbot* que realiza atividades de pesquisa, elaboração e revisão de textos – implementaram os objetivos expostos acima: criaram uma IA que competisse com o *Chat GPT*, a *Claude*, que é regulada pela Carta Universal dos Direitos Humanos da Organização das Nações Unidas. Ao invés de evitar tópicos sensíveis,

ela expõe motivos pelos quais seus princípios a impedem de exteriorizar comportamentos perigosos ou estereotipados. Robinson, CEO da empresa da *Claude* explica:

Usamos Claude para avaliar partes específicas de um contrato e sugerir uma nova linguagem alternativa que seja mais amigável para nossos clientes. Descobrimos que Claude é realmente bom em entender a linguagem — inclusive em domínios técnicos como a linguagem jurídica. Também é muito confiante em redigir, resumir, traduzir e explicar conceitos complexos em termos simples (ROBINSON, 2023, n. p.).

Ainda conforme Robinson (2023), a *Anthropic*, empresa responsável pelo *chatbot*, programou *Claude* com base em dez princípios de cunho constitucional e, por isso, daria respostas mais consistentes e éticas, deixando de inventar fatos ou dar ideias discriminatórias ou ilegais. Embora não seja uma inteligência artificial aberta ao público, como é o *Chat GPT*, *Claude* propõe um avanço à regulamentação de *chatbots* e à eliminação de vieses discriminatórios na programação de algoritmos.

### 5.1.1. Ética algorítmica

Pela ausência de uma concepção de justiça universal para aplicação aos modelos algorítmicos, faz-se necessária a discussão sobre a ética em sua implementação. A ética aplicada à atividade tecnológica implica a análise da implementação dos algoritmos, pelo envolvimento de atores, tanto humanos quanto autômatos, que tomam decisões (ou cálculos), seja por uma função cognitiva ou por uma configuração em seu sistema. Ao considerar os valores e obrigações de uma conduta, tendo a ética como base, deve-se questionar acerca dos limites da tomada de decisões de uma máquina. Apesar de não ser a temática central desta pesquisa, os comportamentos reproduzidos – e as respectivas consequências – pelas máquinas devem ser responsabilizados. Entretanto, é de se indagar quem se responsabilizaria, não sendo a máquina um ser capaz de se sujeitar às penalidades impostas às pessoas.

A impossibilidade de responsabilização da máquina ou do sistema, em si, não pode ocasionar o desrespeito aos direitos e deveres que todos os cidadãos se submetem, em razão de ter tido um autor a tal código, um programador responsável pela configuração dos parâmetros que permitiram eventual conduta a ser responsabilizada, ou ainda, a própria empresa que liberou a utilização de uma inteligência artificial que não estava preparada para efetuar sua finalidade.

Essa questão se torna ainda mais complexa quando se analisa a intenção do agente, se ele tinha algum controle sobre a ação realizada pelo *software* ou se era possível prevê-la, considerando que nem sempre há esta possibilidade de previsão do resultado por parte do programador, ou do fabricante, banir o lançamento do algoritmo nem de o usuário imaginar que sua fala se sucederia em uma discriminação algorítmica.

Desse modo, baseando-se em Rossetti e Angeluci (2021), são deduzidos seis problemas éticos algorítmicos:

- a) evidências inconclusivas;
- b) evidências inextricáveis;
- c) evidências mal direcionadas;
- d) resultados injustos;
- e) efeitos transformativos; e
- f) rastreabilidade.

Quanto às características internas do algoritmo, é possível a ocorrência de erros em sua elaboração que o inutilizem, sendo estas evidências inextricáveis. Estas podem ser demonstradas pelo caso anteriormente mencionado, que está sendo combatido pela Liga da Justiça Algorítmica, em que mulheres negras estão sendo reconhecidas como homens, já que não foram configuradas com dados suficientes para essa diferenciação e, por isso, não haveria utilidade na máquina que apresente tal falha.

Em relação às evidências logicamente incorretas, são inconclusivas ou mal direcionadas, quando não há uma correlação entre causa e efeito. Exemplo de direcionamento equivocado é o software COMPAS, que, programado com dados específicos, passou a atribuir altos níveis de reincidência e periculosidade àqueles detentores desses dados. O que se percebe é que a evidência, nesse caso, poderia ter uma finalidade diversa, isto é, ao invés de se analisar o histórico familiar e antecedentes com álcool e entorpecentes para condenar um indivíduo, poderia se fazê-lo a fim de concretizar a justiça algorítmica, com a implementação de políticas públicas nesse sentido. Diante disso, embora as evidências tenham sido mal direcionadas, foram conclusivas: com os dados disponíveis, as máquinas realizam um cálculo, equivalente a uma decisão, que foi incorreta, mas que havia um vínculo de causalidade presente.

As relações do algoritmo com o contexto social são o foco de crítica desse trabalho, especialmente no que tange o contato do usuário com o algoritmo. Um exemplo já mencionado é o caso da IA *Tay*, que reproduziu comportamentos racistas como reflexo do que era exposto pelos usuários, o que revela os resultados injustos de uma programação algorítmica ao discriminar pessoas por questões raciais e os efeitos transformativos, por ter sido transformada em um porta-voz de racismo pela interação que teve com os usuários, em detrimento da finalidade à qual foi configurada.

A rastreabilidade, enquanto problema ético algorítmico, diz respeito à indisponibilidade de acesso à criação do algoritmo para o usuário, que fica sob controle único de seu programador e da empresa responsável. Essa medida é justificada por uma questão de privacidade de dados, em que os grandes conglomerados alegam que expor tais dados ao público permitiria que fossem apropriados por seus concorrentes, resultando em uma perda de lucro da empresa desenvolvedora, além de acreditarem colocar em risco a segurança das configurações algorítmicas, que poderiam ser utilizados em meios ilícitos e atentatórios à autonomia da companhia. Logo, a impossibilidade da rastreabilidade do algoritmo prejudica a responsabilização pelo comportamento da máquina, visto que, se não é possível determinar quem programou, de que forma ou com quais dados, não é possível atribuir uma penalização a nenhum sujeito.

O conflito ético em questão traz um embate entre a privacidade e a autonomia de uma empresa e a transparência dos algoritmos. A transparência é necessária para inibir eventuais vieses discriminatórios, especialmente no que se trata de uma regulamentação e fiscalização por autoridade governamental, e também para evitar um cenário de insegurança para usuários que se veem sujeitos e numa posição de vulnerabilidade a máquinas e IAs das quais desconhecem a finalidade ou funcionamento. Por esse motivo, deve-se estabelecer um equilíbrio entre os princípios mencionados, que assegure a transparência algorítmica e tranquilize as empresas acerca da análise e utilização dos dados por elas divulgados.

Nos Estados Unidos, em âmbito estadual e municipal, a atenção à transparência algorítmica foi expressa por meio da aprovação de leis nesse sentido. Conforme Pintarelli (2022), no estado da Califórnia, em 2022, a garantia de clareza acerca dos dados utilizados em serviços online voltados às crianças e aos adolescentes foi estabelecida com a aprovação de lei

que restringiu a coleta de seus dados. Esse dispositivo jurídico indica um precedente para os demais países, dado que, ao visar a proteção desse grupo social, é primordial o conhecimento dos algoritmos aos quais ele está exposto. Isto poderia se expressar, no caso do Brasil, como uma alteração legislativa no Estatuto da Criança e do Adolescente para a inclusão de dispositivos sobre a transparência algorítmica e os serviços de internet disponíveis aos infantes.

Também em 2021, foi aprovada, na cidade de Nova York, a lei nº 1894-A que protege os trabalhadores do preconceito algorítmico no recrutamento profissional por processos automatizados. Essa garantia está ligada à representação das pessoas nos dados que regem esses processos automatizados, retornando, então, à discussão de quem seriam os programadores dessas IAs, onde estariam concentrados e quais dados foram utilizados para a configuração de tal sistema, tendo em vista que, no caso em questão, ocorria uma preferência dos recursos humanos por homens caucasianos, em detrimento de outras etnias e gêneros.

## 5.2. Legislação brasileira

O ordenamento jurídico brasileiro condena todas as formas de discriminação que atinjam a vida de qualquer pessoa, direta ou indiretamente. Dessa forma, legislações como a Lei Geral de Proteção de Dados (LGPD), que trata sobre a proteção de dados pessoais e liberdades fundamentais é de suma importância para o desenvolvimento responsável dessa área tecnológica no país. Considerando que a LGPD expressa o princípio da não discriminação no tratamento de dados pessoais, é uma manifestação legislativa que assegura o tratamento igualitário entre pessoas no que diz respeito a seus dados. Todavia, tem-se mostrado insuficiente para regular todos os assuntos suscitados a partir da inteligência artificial e seu consequente racismo algorítmico, resultando numa necessidade de inclusão de dispositivos antidiscriminatórios sobre IA e proteção de dados na legislação, de forma mais específica, à programação e configuração dos algoritmos.

Atualmente, o Projeto de Lei 21/2020 é o projeto que maior abrange a regulação do desenvolvimento de tecnologias de IA no Brasil, estando em tramitação conjunta ao PL 5051/2019 e ao PL 872/2021 por se tratar da mesma matéria. O PL apresenta uma lista de aspectos que requerem regulamentação e atribui a responsabilidade dessas regulamentações a órgãos e entidades setoriais. Além disso, o texto destaca os princípios fundamentais para a IA,

incluindo sua finalidade benéfica, centralidade do ser humano, não discriminação, busca pela neutralidade, transparência, segurança e prevenção, inovação responsável e disponibilidade de dados. Também são apresentados fundamentos para o desenvolvimento e uso da tecnologia, como a autorregulação, a livre manifestação de pensamento e expressão e a proteção de dados pessoais. A maior discussão tem sido gerada em torno de não ter sido inserido o relatório de impacto de sistemas de IA (RIIA) na qual visa estruturar os proventos públicos de aplicação de IA – confiança, segurança e transparência; valores relacionados ao desenvolvimento; identificar e limitar riscos. O enfoque do RIIA é considerar as consequências que cada sistema terá sobre as pessoas, tornando-se um meio de extrema importância para prevenir ou reduzir possíveis danos, como os vieses abordados. Dessa forma, a elaboração da RIIA faz com que os efeitos e os riscos sejam analisados em uma fase pré-implementação de uma nova IA evitando que perpetue os vieses discriminatórios.

A preocupação acerca da regulamentação da inteligência artificial é uma mazela já detectada pelo Legislativo e Judiciário brasileiros, tendo resultado na criação de uma comissão de dezoito juristas para que discutissem a temática, como está no Relatório Final da Comissão de juristas (SENADO, 2022). A comissão de juristas responsável por subsidiar a elaboração de substitutivo sobre inteligência artificial elaborou um extenso relatório – disponibilizado no site do Senado Federal –, com em torno de novecentas páginas, no qual se defendeu, dentre outras questões, em relação ao PL 21/2020, a instituição de uma autoridade competente para fiscalizar o cumprimento das normativas sobre a aplicação e desenvolvimento da IA no Brasil.

Entregue em dezembro de 2022 ao presidente do Senado, o relatório resguardou:

I – a centralidade da pessoa humana; II – o respeito aos direitos humanos e aos valores democráticos; (...) V – a igualdade, a não discriminação, a pluralidade e o respeito aos direitos trabalhistas; VI – o desenvolvimento tecnológico e a inovação; VII – a livre iniciativa, a livre concorrência e a defesa do consumidor; VIII – a privacidade, a proteção de dados e a autodeterminação informativa; IX – a promoção da pesquisa e do desenvolvimento com a finalidade de estimular a inovação nos setores produtivos e no poder público; X – o acesso à informação e à educação, bem como a conscientização sobre os sistemas de inteligência artificial e suas aplicações (SENADO FEDERAL, 2022, p. 19).

Para a elaboração do documento, foram consultados membros da sociedade civil, da academia, do setor privado e do governo, a fim de contribuir com a temática, por meio da

*Ciências Sociais Aplicadas em Revista*, v. 25, n. 45, p. 225-246, Edição especial, 2023.

realização de audiências públicas, divididas por eixos temáticos. Assim, demonstrou-se uma análise combinada entre dispositivos legais, que asseguram a proteção de direitos e da dignidade da pessoa humana, bem como da garantia à inovação tecnológica, segura e harmônica aos preceitos defendidos pela Constituição Federal.

## **6 Procedimentos metodológicos**

Para investigar um tema que está sob constante evolução e tão pouco trabalhado na esfera jurídica, exige do pesquisador uma visão transdisciplinar, não ficando restrito a somente um campo de conhecimento. Adotou-se uma abordagem qualitativa que permitiu aprofundar o entendimento sobre a ética aplicada à tecnologia, à justiça algorítmica, aos vieses discriminatórios, ao racismo algorítmico e à inteligência artificial. Ademais, utilizou-se o método analítico-descritivo como base para a pesquisa, por meio de uma revisão de literatura, por procedimento exploratório e bibliográfico, no qual ocorreu a análise de fontes pertinentes ao tema, a partir de artigos, livros, revistas científicas e a legislação nacional pertinente.

Durante a análise dos materiais, adotou-se uma postura crítica e reflexiva, levando em consideração as diversas perspectivas e debates em torno do uso, aproveitamento, ética e vieses na IA. Buscou-se compreender as implicações sociais, culturais e políticas desses fenômenos, bem como as possíveis consequências para a sociedade. O uso de fontes bibliográficas foi um importante meio para que se fosse possível garantir a validade e a veracidade dos resultados obtidos, pois a partir das fontes se pode analisar as convergências e divergências de opiniões e estudos, engrandecendo a compreensão do tema.

Salienta-se que a abordagem qualitativa e o método indutivo utilizado têm suas limitações. Por mais que se tenha explorado diversos materiais, a pesquisa foi baseada em informações existentes e disponíveis até a data da finalização do artigo. Dessa forma, é válido explicitar que novas pesquisas e avanços podem acrescentar demais *insights* atualizados sobre o tema.

## **7 Resultados e discussões**

A partir da gravidade da problemática exposta, torna-se primordial elencar meios pelos quais é possível mitigar esses vieses dos algoritmos, seja por uma transparência algorítmica,

*Ciências Sociais Aplicadas em Revista, v. 25, n. 45, p. 225-246, Edição especial, 2023.*

pela maior representatividade no ramo de trabalho tecnológico, que sustentaria, então, outras perspectivas, que não seriam mais discriminatórias, e também pela regulamentação do desenvolvimento e produção das inteligências artificiais, orientando-se pelo respeito aos direitos fundamentais, à pluralidade e igualdade de tratamento.

A *International Business Machines Corporation* (IBM), empresa responsável pelo desenvolvimento de soluções de Tecnologia da Informação, destaca a importância de abordar os conteúdos de responsabilidade, transparência, justiça e segurança para aproveitar plenamente os benefícios oferecidos pela IA. Para isso, a IBM (2021) solicita que os governos implementem cinco prioridades para reforçar a adoção de testes, avaliações e estratégias de mitigação de vieses em IA.

Essas prioridades incluem reforçar o conhecimento e a alfabetização em IA para a sociedade, investindo em educação e pesquisas e incluindo a IA nos currículos para que possa haver uma diversificação no desenvolvimento das IAs. Além disso, é essencial exigir avaliações e testes para sistemas de IA de alto risco, garantindo segurança e proteção aos seus consumidores. Dessa forma, cabe ao Estado arcar com os custos que são gerados devido ao investimento na sociedade, já que este deve garantir que as IAs sejam desenvolvidas e aplicadas de forma ética, sem vieses, e que se leve em conta a segurança da população, requerendo, dessa forma, os investimentos em pesquisas e regulamentações apropriadas para lidar com esse sistema. O investimento público nessa área tende a impulsionar a economia, o desenvolvimento social e todos os setores que são tangentes a este; como o financiamento o Estado incentiva a inovação, atrai investimentos de empresas e cria empregos qualificados.

A transparência na IA deve ser promovida por meio da divulgação, informando aos usuários quando o sistema está sendo utilizado, especialmente em casos de tomada de decisões automatizadas pela IA. Os sistemas devem oferecer mecanismos para análise e *feedback* do consumidor, disponibilizando canais de comunicação para esclarecer dúvidas e lidar com reclamações.

Torna-se impossível a detecção do problema sem se ter conhecimento da forma pela qual o algoritmo funciona sem que haja a disponibilidade da configuração do sistema. Por esse motivo, a transparência algorítmica deve ser incorporada como um direito fundamental dos cidadãos, sendo expresso no rol das garantias fundamentais do art. 5º da Constituição Federal,

*Ciências Sociais Aplicadas em Revista*, v. 25, n. 45, p. 225-246, Edição especial, 2023.

a partir de uma emenda constitucional, tendo em vista a relevância da temática, que não mais pode ser contemplada em um projeto de lei. Apenas assim será possível analisar a configuração dos algoritmos para que se previna a discriminação de indivíduos, com base em vieses. A regulamentação das IAs no texto constitucional brasileiro seria uma medida inovadora no contexto mundial, seguindo o exemplo chileno, garantindo a segurança dos cidadãos diante de um contexto de vulnerabilidade perante os softwares, podendo, para isso, se utilizar dos resultados do relatório da comissão de juristas sobre inteligência artificial, a fim de concretizar preceitos de igualdade, não discriminação, privacidade e proteção de dados.

Por fim, é fundamental estabelecer limitações universais para o uso de IA e adotar práticas de licenciamento responsáveis, proibindo a utilização da IA como meio de vigilância, bem como qualquer forma de discriminação e infração aos direitos humanos. Essas medidas visam garantir um desenvolvimento ético e responsável da IA protegendo os interesses e valores da sociedade como todo.

## **8 Considerações finais**

A Inteligência Artificial está presente direta e indiretamente na vida de toda a população e em todos os lugares e ocupa uma importância cada vez mais crescente. A IA utiliza algoritmos para sua execução e reproduz comportamentos humanos por meio das máquinas. No entanto, a qualidade dos dados que são inseridos nos sistemas de IA e os vieses dos algoritmos podem levar à discriminação algorítmica que poderá surgir devido à qualidade dos dados, restrições legais, limitações algorítmicas, contexto inadequado e interpretação incorreta.

O enviesamento dos algoritmos reflete o contexto social e as opiniões dos programadores, que podem ser permeadas por discriminações e exclusões. Dentre os exemplos citados pode-se perceber como exemplos de discriminação algorítmica o *Beauty.ai*, concurso de beleza em que as IAs escolheram majoritariamente pessoas brancas como as mais bonitas, sendo as vencedoras do concurso, ou nos casos de sistemas de IA que são utilizados nos tribunais dos EUA na qual categorizaram pessoas negras como prováveis reincidentes de criminalidade, enquanto as pessoas brancas dificilmente eram categorizadas como tal. Dessa forma, para que se possa haver a mitigação da discriminação algorítmica, é necessário que se garanta a qualidade dos dados, o contexto adequado e a interpretação correta, além de que

ocorra uma conscientização da sociedade sobre o problema. O trabalho conjunto entre criadores e desenvolvedores das IAs com os legisladores é de extrema importância, pois assim podem conjuntamente garantir o desenvolvimento ético e responsável da IA.

Por fim, afirma-se que os vieses algorítmicos possuem grande impacto na sociedade, como por exemplo, o encarceramento em massa de pessoas negras devido a um sistema de reconhecimento facial, na qual, conseqüentemente reforça estereótipos e continua a perpetuar o racismo estrutural. Além disso, com a falta de transparência e legislação, gera uma erosão da confiança na tecnologia em que resultam em resistência e rejeição perante a sociedade, o que, de certo modo, gera um bloqueio no desenvolvimento dessa área; não há uma responsabilização decente para que a população se sinta segura em sua utilização. A ética algorítmica exerce um papel de tamanha importância pois questiona os limites das respostas automatizadas dos sistemas e a responsabilização dos indivíduos envolvidos na implementação dos algoritmos. A busca por soluções éticas e regulamentações adequadas pode ajudar na mitigação do impacto social dos vieses algorítmicos.

## Referências

ABREU, A. J. A.; FURTADO, K. C. S.; SANTOS, R. K. C. Inteligência artificial e preconceito de identidade de gênero: o problema do viés na construção das IA's e a perpetuação das discriminações em sociedades previamente discriminatórias. **COR LGBTQIA+**, [S. l.], v. 1, n. 3, p. 229–247, 2022. Disponível em: <https://revistas.ceeinter.com.br/CORLGBTI/article/view/551>. Acesso em: 25 mai. 2023.

BRASIL. Câmara dos Deputados. **Projeto de Lei nº 21/2020**. Estabelece fundamentos, princípios e diretrizes para o desenvolvimento e a aplicação da inteligência artificial no Brasil; e dá outras providências. Brasília, DF: Câmara dos Deputados, 2020. Disponível em: <https://www.camara.leg.br/propostas-legislativas/2236340>. Acesso em: 31 mai. 2023.

BRASIL. [Constituição (1988)]. **Constituição da República Federativa do Brasil de 1988**. Brasília, DF: Presidência da República, [2020]. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/Constituicao/Constituicao.htm](http://www.planalto.gov.br/ccivil_03/Constituicao/Constituicao.htm). Acesso em: 23 mai. 2023.

BRASIL. Senado Federal. **Projeto de Lei nº 5051/2019**. Estabelece os princípios para o uso da Inteligência Artificial no Brasil. Brasília, DF: Senado Federal, 2019. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/138790>. Acesso em: 31 mai. 2023.

*Ciências Sociais Aplicadas em Revista*, v. 25, n. 45, p. 225-246, Edição especial, 2023.

BRASIL. Senado Federal. **Projeto de Lei nº 872/2021**. Dispõe sobre o uso da Inteligência Artificial. Brasília, DF: Senado Federal, 2019. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/147434>. Acesso em: 31 mai. 2023.

COLAB PUC Minas. **Racismo algorítmico**: o preconceito na programação. Disponível em: <https://blogfca.pucminas.br/colab/racismo-algoritmico-o-preconceito-na-programacao/>. Acesso em: 25 mai. 2023.

DWORKIN, Ronald. **O império do direito**. Tradução: Jefferson Luis Camargo. São Paulo: Martins Fontes, 1999.

FIALHO, Ana Carolina Mendes. Reconhecimento facial: entenda o impacto da inteligência artificial na vida da população negra. In: **Exame**. [S.l.], 2022. Disponível em: <https://exame.com/esg/reconhecimento-facial-entenda-o-impacto-da-inteligencia-artificial-na-vida-da-populacao-negra/>. Acesso em: 30 abr. 2023.

GILLESPIE, Talles. **A relevância dos algoritmos**. Tradução: Amanda Jurno. Revista Parágrafa, São Paulo, v. 6, n. 1, p. 95-121, jan./abr. 2018. Disponível em: [https://edisciplinas.usp.br/pluginfile.php/5971548/mod\\_resource/content/1/722-2195-1-PB.pdf](https://edisciplinas.usp.br/pluginfile.php/5971548/mod_resource/content/1/722-2195-1-PB.pdf). Acesso em: 25 mai. 2023.

Miss Universo julgado por inteligência artificial revela racismo das máquinas. In: **Olhar Digital**. [S.l.], 9 set. 2016. Disponível em: <https://olhardigital.com.br/2016/09/09/seguranca/miss-universo-julgado-por-inteligencia-artificial-revela-racismo-das-maquinas/>. Acesso em: 30 abr. 2023.

Mitigar o preconceito na inteligência artificial. In: **IBM Blog**. [S.l.]. Disponível em: <https://www.ibm.com/blogs/ibm-comunica/mitigar-preconceito-inteligencia-artificial/>. Acesso em: 24 mai. 2023.

OLIVEIRA, Evandro. **Coded Bias e a Liga da Justiça Algorítmica**. Capital Digital, 2022. Disponível em: <https://capitaldigital.com.br/coded-bias-e-a-liga-da-justica-algoritmica/>. Acesso em: 25 mai. 2023.

PINTARELLI, Camila. **Justiça algorítmica em uma perspectiva de gênero**. Jota, 2022. Disponível em: <https://www.jota.info/opiniao-e-analise/artigos/justica-algoritmica-em-uma-perspectiva-de-genero-29102022>. Acesso em: 25 mai. 2023.

PIXEL D. **Conheça Claude**, o chatbot que promete competir com o ChatGPT da OpenAI. Disponível em: <https://pixeld.news/conheca-claude-o-chatbot-que-promete-competir-com-o-chatgpt-da-openai/>. Acesso em: 26 mai. 2023.

RODRIGUES, Fernanda. Relatórios de Impacto e a Mitigação de Vieses em IA. In: **Iris BH**. [S.l.]. Disponível em: <https://irisbh.com.br/relatorios-de-impacto-e-a-mitigacao-de-vieses-em-ia/>. Acesso em: 24 mai. 2023.

ROSSETTI, Regina; ANGELUCI, Alan. Ética Algorítmica: questões e desafios éticos do avanço tecnológico da sociedade da informação. **Galáxia**, São Paulo, v. 1, n. 46, p. 1-18, jul. 2021. Disponível em: <https://revistas.pucsp.br/index.php/galaxia/article/view/50301>. Acesso em: 26 mai. 2023.

SENADO FEDERAL. Comissão de Juristas responsável por subsidiar elaboração de substitutivo sobre inteligência artificial no Brasil. **Relatório Final**. Brasília, DF: Senado Federal, 2022.

SENADO NOTÍCIAS. **Comissão conclui texto sobre regulação da inteligência artificial no Brasil**. Agência Senado, 2022. Disponível em: <https://www12.senado.leg.br/noticias/materias/2022/12/06/comissao-conclui-texto-sobre-regulacao-da-inteligencia-artificial-no-brasil>. Acesso em 26 mai. 2023.

SILVA, Evelyn Melo. Pela proteção dos neurodireitos no Brasil. **Revista Consultor Jurídico**, 2022. Disponível em: <https://www.conjur.com.br/2022-ago-29/evelyn-silva-protecao-neurodireitos>. Acesso em: 30 mai. 2023.

SILVA, Tarcízio. **Racismo Algorítmico**: inteligência artificial e discriminação nas redes digitais. São Paulo: Edições Sesc, 2022.

SOARES, Marcelo Negri; CENTURIÃO, Luís Fernando; TOKUMI, Carine Alfama Lima. Inteligência artificial e discriminação: um panorama sobre a antagonização entre exclusão e o Estado Democrático de Direito Brasileiro à luz dos direitos da personalidade. **Revista Direitos Sociais e Políticas Públicas (UNIFAFIBE)**, v. 10, n. 2, p. 567-597, 2022. Disponível em: <https://portal.unifafibe.com.br/revista/index.php/direitos-sociais-politicas-pub/article/view/1311>. Acesso em: 25 mai. 2023.

SUPREMO TRIBUNAL FEDERAL. Ministro anula condenação baseada apenas em reconhecimento fotográfico do acusado na fase de inquérito. In: **Portal STF**. [S.l.]. Disponível em: <https://portal.stf.jus.br/noticias/verNoticiaDetalhe.asp?idConteudo=418535>. Acesso em: 31 mai. 2023.

TACCA, Adriano; ROCHA, Leonel Severo. Inteligência artificial: reflexos no sistema do Direito. **Revista do Programa de Pós-Graduação em Direito da UFC**: v. 38.2, p. 53-68, jul./dez. 2018. Disponível em: <http://periodicos.ufc.br/nomos/article/view/20493>. Acesso em: 25 mai. 2023.