



DISCRIMINAÇÃO ALGORÍTMICA, VULNERABILIDADE E A PROTEÇÃO AOS DIREITOS FUNDAMENTAIS NA ERA DIGITAL¹

ALGORITHMIC DISCRIMINATION, VULNERABILITY, AND THE PROTECTION OF FUNDAMENTAL RIGHTS IN THE DIGITAL ERA

Mônia Clarissa Hennig Leal²
Lucas Moreschi Paulo³

Resumo: O presente examina a crescente preocupação no campo jurídico em relação à discriminação algorítmica e à proteção dos direitos fundamentais na era digital. Utilizando uma abordagem lógico-dedutiva, a pesquisa analisa as implicações sociojurídicas complexas da automação de decisões por algoritmos, com foco na perpetuação de desigualdades e injustiças, especialmente contra grupos vulneráveis. A partir disso, o estudo propõe o estabelecimento de um panorama acerca do problema, a causa imediata e o campo de possibilidades jurídico-constitucionais a enfrentar a questão da prejudicialidade da discriminação algorítmica à população socialmente marginalizada. Conclui-se que a compreensão dos grupos vulneráveis, a análise da discriminação algorítmica e a eficácia dos direitos fundamentais são elementos essenciais para abordar a questão no contexto do constitucionalismo digital.

Palavras-chave: Algoritmos, Discriminação, Direitos fundamentais.

¹ Este artigo é resultante das atividades do projeto de pesquisa “Teoria da essencialidade” (*Wesentlichkeitstheorie*) e discriminação algorítmica: *standards* protetivos em face do Supremo Tribunal Federal e da Corte IDH – proposta de parâmetros de controle”, financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq (Bolsa de Produtividade em Pesquisa – Processo 309115/2021-3). A pesquisa é vinculada ao Grupo de Pesquisa “Jurisdição Constitucional aberta” (CNPq) e desenvolvida junto ao Centro Integrado de Estudos e Pesquisas em Políticas Públicas – CIEPPP e ao Observatório da Jurisdição Constitucional Latino-Americana (ambos financiados pelo FINEP e ligados ao Programa de Pós-Graduação em Direito – Mestrado e Doutorado da Universidade de Santa Cruz do Sul – UNISC). Também se insere no âmbito do projeto de cooperação internacional “Observatório da Jurisdição Constitucional Latino-Americana: recepção da jurisprudência da Corte Interamericana de Direitos Humanos e sua utilização como parâmetro para o controle jurisdicional de Políticas Públicas pelos Tribunais Constitucionais”, financiado pela Capes (Edital PGC/02/2015 – Processo 88881.1375114/2017-1 e Processo 88887.137513/2017-00).

² Com Pós-Doutorado na Ruprecht-KarlsUniversität Heidelberg (Alemanha) e Doutorado em Direito pela Universidade do Vale do Rio dos Sinos – Unisinos (com pesquisas realizadas junto à Ruprecht-KarlsUniversität Heidelberg, na Alemanha). Professora do Programa de Pós-Graduação em Direito – Mestrado e Doutorado da Universidade de Santa Cruz do Sul – UNISC, onde ministra as disciplinas de Jurisdição Constitucional e de Controle Jurisdicional de Políticas Públicas, respectivamente. Coordenadora do Grupo de Pesquisa “Jurisdição Constitucional aberta”, vinculado ao CNPq. Bolsista de produtividade em pesquisa do CNPq. Lattes: <<http://lattes.cnpq.br/6628165246247243>>. Orcid: <<https://orcid.org/0000-0002-3446-1302>>. E-mail: <moniah@unisc.br>.

³ Advogado. Doutorando em Direito no Programa de Pós-Graduação em Direito – Mestrado e Doutorado da Universidade de Santa Cruz do Sul (UNISC), bolsista do Programa de Suporte à Pós-Graduação de Instituições Comunitárias de Educação Superior (PROSUC) da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Mestre e graduado em Direito pela Fundação Escola Superior do Ministério Público (FMP), foi bolsista institucional do PPGD da FMP. Integrante do Grupo de “Pesquisa Jurisdição Constitucional Aberta”, coordenado pela Prof.^a Dr.^a Mônia Clarissa Hennig Leal, vinculado ao PPGD – Mestrado e Doutorado da UNISC, financiado pelo CNPq. Pesquisador do Grupo de Pesquisa Colisão de Direitos Fundamentais e o Direito como Argumentação, coordenado pelo Prof. Dr. Anizio Pires Gavião Filho, do PPGD – Mestrado da FMP e vinculados no CNPq ao Grupo de Estudos Tutelas à Efetivação dos Direitos Transindividuais. Membro da *Argumentation Network of the Americas* - ANA. Currículo Lattes: <http://lattes.cnpq.br/4330914363996350>. ORCID: <https://orcid.org/0000-0003-4583-4853>. E-mail: lucasmoreschipaulo@gmail.com.



Abstract: The present examines the growing concern in the legal field regarding algorithmic discrimination and the protection of fundamental rights in the digital era. Using a logical-deductive approach, the research analyzes the complex socio-legal implications of automated decisions made by algorithms, with a focus on perpetuating inequalities and injustices, particularly against vulnerable groups. Based on this analysis, the study proposes an overview of the problem, its immediate causes, and the legal-constitutional possibilities to address and mitigate the harmful effects of algorithmic discrimination. It concludes that understanding vulnerable groups, examining algorithmic discrimination, and ensuring the efficacy of fundamental rights are essential elements for addressing this issue within the context of digital constitutionalism.

Keywords: Algorithm, Discrimination, Fundamental Rights.

1 Introdução

Na era digital, a ascensão da automação de tomada de decisões, por meio de algoritmos, trouxe à tona uma nova forma de discriminação - a algorítmica. Essa prática pode inadvertidamente ou intencionalmente perpetuar desigualdades e injustiças humanas, frequentemente associadas a estereótipos desabonadores. Como resultado, questões sobre as implicações sociojurídicas da discriminação algorítmica têm ganhado destaque e demandado uma análise aprofundada por diversos setores, dentre os quais, está o jurídico.

A discriminação algorítmica, uma questão emergente e de grande importância no mundo contemporâneo, permeia todos os aspectos de nossas vidas. A tecnologia, que é uma ferramenta poderosa para o progresso, também pode ser uma fonte de desigualdade e injustiça quando mal utilizada ou mal compreendida em seus processos e resultados. Este estudo busca colocar mais um tijolo na busca da construção semântica e epistêmica acerca tanto da compreensão do fenômeno da discriminação algorítmica quanto das respostas atribuíveis pelo direito, seja pela regulação ou pela jurisprudência da jurisdição constitucional.

A discriminação algorítmica representa um desafio complexo que transcende fronteiras disciplinares, e, à medida que algoritmos e inteligência artificial (“IA”) desempenham papéis cada vez mais proeminentes em nossa sociedade, arvora-se com mais gravidade a necessidade de um olhar atento à proteção dos direitos fundamentais. Para tanto, o grau de previsibilidade – fruto do esforço de técnicos e juristas – não deve ser subestimado, devendo alcançar um conhecimento acerca das fontes e das razões para a ocorrência de resultados discriminatórios pelas aplicações algorítmicas, que – adianta-se – são diversas e complexas, envolvendo desde a falta de diversidade na equipe de desenvolvimento, até a tendência dos algoritmos de perpetuar os preconceitos existentes na sociedade a partir da construção desse raciocínio autonomamente.



Através de um método lógico-dedutivo, o objetivo é delinear um panorama acerca do problema (discriminação), a causa imediata (algoritmos) e o campo de possibilidades jurídico-constitucionais (direitos fundamentais) a enfrentar a questão. Para tanto, primeiro, se buscará a compreensão dos grupos vulneráveis e da discriminação estrutural, explorando a problemática da discriminação algorítmica e, por fim, buscando construir uma noção geral de eficácia dos direitos fundamentais em casos de discriminação automatizada no novo contexto do constitucionalismo digital que, não poderá ter depositada sua expectativa nos modelos regulatórios que vierem a ser criados, mas na eficácia dos direitos em si.

2. Vulnerabilidade, discriminação estrutural e compensação pelo direito

A compreensão das vulnerabilidades, especialmente de grupos estruturalmente discriminados, é essencial para traçar um cenário minimamente fidedigno aos desafios que a nova realidade oferece à comunidade e ao direito. A discriminação estrutural, muitas vezes disfarçada nas sombras da sociedade, afeta profundamente pessoas e grupos, criando um fosso entre as possibilidades de acesso a direitos fundamentais e o seu efetivo exercício. Assim, é necessário estabelecer critérios e noções a partir do direito acerca da vulnerabilidade, da discriminação estrutural e dos grupos vulneráveis, bem como compreender como o direito pode oferecer instrumentos de compensação dessa desigualdade.

O conceito de vulnerabilidade no direito está em sua raiz principiológica relacionado ao princípio da igualdade, do qual também se é derivado. A Constituição Federal estabelece o direito à igualdade geral no *caput* do artigo 5º, garantindo a todos que estejam no país a inviolabilidade de sua igualdade. A igualdade, portanto, não é apenas um direito fundamental, mas sim um valor estruturante do próprio compromisso democrático-constitucional estabelecido na Carta da República, que também prevê a construção de uma sociedade livre, justa e solidária. Livre para se desenvolver, justa para reconhecer a dignidade em si e nos outros, atribuindo-lhes o que lhes é de direito, e solidária e fraternal, para pertencer ao sentimento de pátria.

A igualdade enquanto valor transcende sua origem como revolucionário-liberal e direito fundamental de primeira geração, fundamentando também os direitos de segunda geração, que visam à concretização da igualdade material, indo além da mera igualdade formal perante a lei, englobando aspectos econômicos, sociais e culturais. O direito à igualdade é um princípio jurídico fundamental que requer a aplicação uniforme da lei para garantir um tratamento equitativo, evitando discriminações indevidas (Alexy, 2015, p. 396). Embora o direito possa



aliviar os efeitos de desigualdades fáticas, não pode suprir diferenças inerentes como saúde ou beleza. A igualdade, conforme Alexy (2015, p. 398), implica a aplicação das mesmas normas a todos que se encontram sob a mesma hipótese fática, independente das características pessoais. A igualdade material, então, busca nivelar os desiguais, fornecendo um tratamento diferenciado para vulneráveis, não para privilegiá-los, mas para equipará-los a outros em posição mais favorável, uma abordagem sustentada pela Constituição brasileira como uma cláusula democrática, não existindo desigualdades iguais, ou desigualdades fáticas universais, por isso que a ordem jurídica deve se preocupar tanto na valoração da igualdade quanto na possibilidade de diferenciação do tratamento (Alexy, 2015, p. 400). A igualdade material busca equilibrar as condições sociais, oferecendo a indivíduos desiguais ou vulneráveis um tratamento jurídico diferenciado, não como um privilégio, mas para assegurar igualdade de condições. Esse princípio, presente no texto constitucional brasileiro, visa proteger os vulneráveis em todas as suas interações, promovendo um tratamento equitativo frente aos não vulneráveis, ou aos menos vulneráveis em casos de múltiplas vulnerabilidades coexistentes.

A Constituição Federal direciona o ordenamento jurídico na proteção de grupos vulneráveis específicos, como consumidores, idosos e crianças e adolescentes, através de legislação dedicada. A interpretação e aplicação de todas as normas devem considerar a vulnerabilidade como um critério essencial para prevenir danos injustos – critérios de diferenciação e finalidade de distinção para com congruência do critério e razão do fim desejado (Ávila, 2016, p. 192-194). Essa perspectiva é reforçada pela noção de que a igualdade permite tratamento desigual apenas para corrigir desigualdades, rejeitando discriminações arbitrárias que não se justifiquem sob critérios racionais e objetivos (Alexy, 2015, p. 409). A vulnerabilidade é escopo interpretativo da ordem jurídica, impondo um ônus de maior cuidado pelo intérprete-aplicador, sendo tarefa sua demonstrar argumentativamente o grau de desigualdade e atribuir o critério da vulnerabilidade comparativa, bem como a consequência jurídica ou mecanismo de mitigação. Deste modo, se não houver razão suficiente para permitir dispensar um tratamento igualitário, o tratamento desigual (mais protetivo) é obrigatório na medida das desigualdades e da vulnerabilidade constatada.

Eis o cerne do reequilíbrio social que o Estado Democrático oferece. Há, contudo, que se ir um pouco mais além na busca conceitual do estado da arte da vulnerabilidade, eis que a desigualdade pode ainda figurar-se a partir de uma forte interseccionalidade que impulsiona a lesividade da condição, como é o caso de pessoas que se enquadram em múltiplas categorias de vulnerabilidade e são objetificadas, até mesmo inconscientemente, como menores em



importância e dignidade pelas estruturas sociais, ou então de modo intencional (Moreira, 2019, p. 64). Há, ainda, a discriminação estrutural que, muitas vezes, traz em si não apenas o conceito de interseccionalidade, mas também trabalha com a categoria das minorias – enquanto grupo de pessoas que se encontra em posição de desvantagem em relação à sociedade como um todo, sejam por questões econômicas, políticas, de origem étnica, nacional ou linguística, orientação sexual, escolhas religiosas e outras, bem como sejam elas em número reduzido ou não dentro da sociedade observada (Carbonell, 2000).

Não se trabalhará a partir do recorte das minorias, mas sim dos grupos vulneráveis – mais adiante grupos vulneráveis em relação ao uso dos algoritmos –, cujos indivíduos podem pertencer, ou não, a uma minoria e, daí, demandarem atenção ainda mais tutelar do direito. A discriminação sofrida por esses, entretanto, é de natureza estrutural, e essa demanda não apenas um olhar atento em termos conceituais, mas em termos da prática da jurisdição e do dever de proteção estatal. Essa discriminação frequentemente resulta em desigualdades na cidadania e direitos individuais, agindo de forma sistêmica, perpetuando desigualdades e injustiças a partir de estereótipos estratificantes, ocorrendo quando da invisibilidade social ou quando as regras, práticas, políticas e normas em uma sociedade beneficiam ou prejudicam grupos de pessoas de forma consistente e duradoura, enraizadas na rotina da sociedade, resultando em disparidades significativas em termos de fruição efetiva da cidadania e das garantias individuais, criando desvantagens significativas para os prejudicados, e ainda maior para grupos vulneráveis que sofrem ainda mais com a discriminação estrutural, experimentando desvantagens ainda mais significativas.

Como se percebe o termo “discriminação estrutural” carrega em si uma carga simbólica de difícil definição, o que demonstra sobretudo as possibilidades de sua caracterização no mundo fenomênico. As discriminações estruturais sofridas por grupos vulneráveis podem se dar por muitos elementos culturais e históricos, da própria formatação do *status* social em certa sociedade, como pode significar situações cujo nascedouro sejam mais contemporâneas, implicando igualmente em gravosa discriminação estrutural. Leal e Alves (2023, p. 8) trazem importantes elementos para a identificação de uma discriminação estrutural, sem exigir a presença de todos eles cumulativamente, ou a eles se limitar. Vejamos:

(I) A discriminação estrutural exige a presença de um grupo ou grupos de pessoas, tratando-se de uma situação coletiva. O grupo deve apresentar (a) características imutáveis ou imodificáveis por vontade própria (b) ou que estejam relacionadas a fatores históricos de práticas discriminatórias; (c) e que este grupo seja de minorias ou maiorias.



(II) Que esse grupo se encontre em uma situação sistemática e histórica de exclusão, marginalização ou subordinação que lhes impeça de alcançar as condições básicas de desenvolvimento humano. Este é o núcleo central da noção de discriminação estrutural: constatar uma situação generalizada de desvantagem, exclusão, subordinação, marginalização ou submissão como obstáculo ao gozo dos direitos fundamentais.

(III) Que a situação de exclusão, marginalização ou subordinação se centre em uma zona geográfica determinada, ou generalizado em todo o território do Estado, podendo ser ainda uma situação intergeracional.

(IV) Que a discriminação estrutural afete grupos de vítimas de discriminação indireta ou de discriminação de fato, por atuações, aplicações de medidas ou ações implementadas pelo Estado. Ou seja, independe da intenção da norma, da neutralidade ou da menção expressa de alguma distinção ou restrição explícita, devendo ser apurado se há violação à igualdade material para determinado grupo vulnerável.

Como se percebe, há uma certa amplitude com conceitual a abarcar situações que de fato passam de vulnerabilidades circunstanciais, de condições naturais ou morais dos indivíduos (ainda que essas sejam “razões” para a prática ou perpetuação de preconceitos), alcançando uma situação de verdadeira estratificação excludente. Assim, o nascedouro da discriminação estrutural passa por problemas associados a vícios na fruição da igualdade, tanto em sua face de princípio norteador de todo o sistema jurídico, quanto em seu aspecto jusfundamental de proteção individual e, nesses termos, ao não tratamento discriminatório. Isto é, à vedação ao tratamento discriminatório não justificado.

Tem-se o seguinte fato: tais valores decorrentes da igualdade nem sempre foram reafirmados de modo efetivo (ao menos não meramente retórico) através dos tempos. Há uma inegável desigualdade material, e de efetivar a igualdade formal, em uma organização jurídico-estatal que, mesmo com ações afirmativas pelo Estado. Há uma sistemática manutenção da discriminação estrutural que cidadãos são subjugados por condições comunitárias específicas (Saba, 2005, p. 126), que tanto “justificam” a posição de desvantagem quanto a retroalimentam e, nesse ponto – como se verá mais adiante – as aplicações algorítmicas são uma verdadeira indústria de reprodução das discriminações estruturais.

Acerca da discriminação estrutural, Sagüés (2018) propõe que a desigualdade sistemática seja atacada por uma proposta de transformação social na instância superior da igualdade estrutural, permitindo que decorra desta concepção a necessária tutela especial a esses grupos, sobretudo das ações afirmativas – ainda que em um necessário caráter excepcional e transitório – pensado para que venha a inexistir com a eficácia das medidas no tempo.

O Estado tem o dever de criar condições para a efetiva proteção dos direitos fundamentais, como a igualdade. Na ausência de ação legislativa ou constitucional adequada, é essencial que o judiciário possa intervir, seja preenchendo lacunas protetivas ou agindo repressivamente contra inconstitucionalidades e ilegalidades, respeitando os limites normativos



existentes. Desse modo, é importante lembrar-se que as medidas de ações afirmativas que consistem em práticas não discriminatórias começam a partir de uma suposição fática fundamental de natureza coletiva: age contra pressupostos de natureza coletiva, um grupo sistematicamente excluído, subjugado ou marginalizado, que podem, mas nem sempre, ser historicamente discriminados – afro-americanos, mulheres, pessoas com deficiência e assim por diante (Sagüés, 2018).

A natureza transindividual e coletiva da discriminação estrutural reflete que a medida de ação afirmativa que implementa mecanismos de não discriminação com real possibilidade de transformação da realidade também recebem elementos coletivos. Trata-se de uma tutela ampla, portanto, não apenas inividual, sendo um desdobramento do dever de proteção amplo, e não uma “mera” postulação individual jusfundamental. O que não significa que a identificação de problemas estruturais, sistemáticos e coletivos não possa ser dada, em um primeiro momento, no viés individual, e depois ser ampliada para a tutela coletiva (Sagüés, 2018).

Os juízos procedimentais, constitucionalmente estabelecidos, sobretudo em ditames e situações notavelmente estruturantes, devem respeitar a concepção de ser, o instrumento processual, uma segunda face de uma mesma moeda na qual estão os direitos. É importante que haja a concretização de direitos fundamentais. É importante, sendo também um dever do Estado, que haja a erradicação da pobreza e a redução da desigualdade (Sagüés, 2018), ou ao menos a busca ativa desses propósitos, e não apenas enquanto um ímpeto apassivador.

Diante das situações de discriminação por vulnerabilidade, pertencimento a grupos vulneráveis ou discriminação estrutural, os poderes da nação devem proteger os indivíduos das agressões injustas aos seus direitos, cabendo ao poder judiciário interpretar os núcleos de significação dos textos normativos provenientes dos demais poderes, bem como tutelar por último a correta, proporcional e adequada resposta aos injustos percebidos. Na realidade da inteligência artificial, há um vácuo normativo acerca da temática, mas as injustas agressões e violações à direitos fundamentais não esperam pela legislação para conhecer o adequado tratamento jurídico, sendo cabível ao poder judiciário dar, em caráter tópico e jurisprudencial, respostas cabíveis e parâmetros racionais de interpretação e aplicação do direito às demandas envolvendo discriminação algorítmica. Considerando a complexidade da vulnerabilidade e da discriminação estrutural, é fundamental levar em conta esses elementos para abordar a discriminação algorítmica, que atinge especialmente os grupos vulneráveis, a partir de interseccionalidades, ou seja, múltiplas origens e motivações injustificadas que trazem severas desproporções entre a eficácia dos direitos fundamentais e sua fruição nas circunstâncias fáticas



do cotidiano. O direito deve reequilibrar proporcional e gradativamente essa balança, trazendo clarões de garantias fundamentais.

3. O fenômeno emergente da discriminação algorítmica

A utilização das Inteligências Artificiais para a tomada de decisões é um feito realizado desde os anos 80, quando empresas do mercado de ações passaram a contar com as ferramentas para conseguir resultados muito acima da concorrência. A automatização da tomada de decisão para comprar, vender, alocar, realocar ou *opt out* demonstrava uma certeza e uma aptidão quase absoluta para os melhores resultados acionários. De lá para cá, as IAs se expandiram, sendo aplicadas em muitos campos da vida cotidiana, como na agricultura, na gestão de consumo de energia, no controle de vazão de água, ou mesmo para solucionar problemas antes impensáveis. Eis um aspecto interessante dos algoritmos: eles são uma cláusula aberta, e sua capacidade dependerá da habilidade que seu programador der à programação inaugural (Steiner, 2012).

Além disso, como bem demonstra Heikkilä (2023b), as máquinas não são conscientes, não tendo dentro de si, ao longo das linhas de sua codificação técnico-genética, qualquer autoimplicância ou autorreflexão acerca de seus atos, das implicações éticas ou das consequências potencialmente nocivas de suas atitudes. Tais preocupações até podem se fazer presentes, mas não a título de consciência, e sim de construção de *feedbacks* e *reports* (relatórios). O que pode ser objeto de desejo por parte de desenvolvedores e usuários, a consciência da IA demandaria, primeiro, conhecer a própria essência da consciência humana, algo ainda desconhecido. E, portanto, não completamente auferível *a priori* ou mesmo implicado em previsibilidade.

Os algoritmos são conjuntos de instruções lógicas que um programa de computador segue para realizar uma tarefa específica de maneira sistemática, seguindo uma série de instruções pré-estabelecidas para realizar uma determinada tarefa. Assim, quando as IAs discriminam automaticamente pessoas no processo de sua tomada de decisão, não agem com consciência acerca da prejudicialidade ou da ilicitude do *outcome* (resultado) (Huckins, 2023). O termo discriminação algorítmica se refere às discriminações promovida pelos algoritmos ao discriminar pessoas com base em características e padrões predeterminados, o que conduz naturalmente a resultados injustos e desiguais. Tais discriminações estão, geralmente, relacionadas com a amostragem de dados coletada e armazenada na base de dados utilizada pelo algoritmo em questão.



Algoritmos têm potencial discriminatório simplesmente por utilizarem dados. Se os dados históricos forem prejudiciais ou não representativos da população, os algoritmos podem refletir esses dados e perpetuar padrões não desejáveis. O problema é que não há inteligência artificial ou algoritmos sem o uso de dados. A qualidade da inteligência e do processamento de dados depende diretamente do volume e da qualidade destes. Fatalmente, os algoritmos acabam discriminando pessoas com base em vieses herdados ou não identificados por sua base de dados ou por seus programadores.

Características como gênero, raça e classe social facilmente servem para classificar pessoas, especialmente em um contexto de discriminação histórica e estrutural, onde estes aspectos estão arraigados na própria lógica de funcionamento da sociedade, e isso tem especial potencial lesivo dentro da linguagem de máquina, que faz automaticamente e de modo muito eficiente a clusterização, isto é, a formação de grupos *ad hoc* não necessariamente autoevidentes em suas correlações, aos olhos humanos, pelo menos. Tais clusterizações geralmente dizem respeito a grupos vulneráveis que também são identificáveis pela sua situação nada privilegiada de serem sujeitos passivos da discriminação estrutural. Isto é, uma discriminação estruturada nas raízes identitárias da sociedade que, muitas vezes, aponta para minorias marginalizadas, ou ainda interseccionalidades que prejudicam o desenvolvimento pessoal, comunitário e social na comunidade. Tais grupos de pessoas são identificados por semelhanças quanto à questões de ordens econômicas, políticas, de origem étnica, nacional ou linguística, orientação sexual, escolhas religiosas e outras (Carbonell, 2000).

Vê-se, portanto, que os grupos vulneráveis têm um especial prejuízo em relação ao uso dos algoritmos, uma vez que não raras vezes a conclusão das aplicações de IA podem apontar a vedações de determinados bens e serviços a determinado *cluster*, ou, então, estereotipar pessoas em tônicas de identidade não acertadas. A discriminação nessa tônica mais gravosa é de natureza estrutural, e essa demanda um olhar atento tanto em termos conceituais, quanto em termos da prática da jurisdição e do dever de proteção estatal.

Segundo Heikkilä (2023a), há ferramentas que permitem avaliar o quão enviesados são os modelos de imagens gerados IA. A utilização de tais modelos de IA por poderem ter efeitos discriminatórios, autorizam a exigência de que tais ferramentas para os avaliar sejam mais amplamente divulgadas e tornadas públicas, o que ainda não é uma realidade⁴. Há como identificar tais padrões na discriminação algorítmica. Os pesquisadores pediram para que a IA

⁴ Uma delas está disponível em <<https://huggingface.co/spaces/society-ethics/StableBias>>.

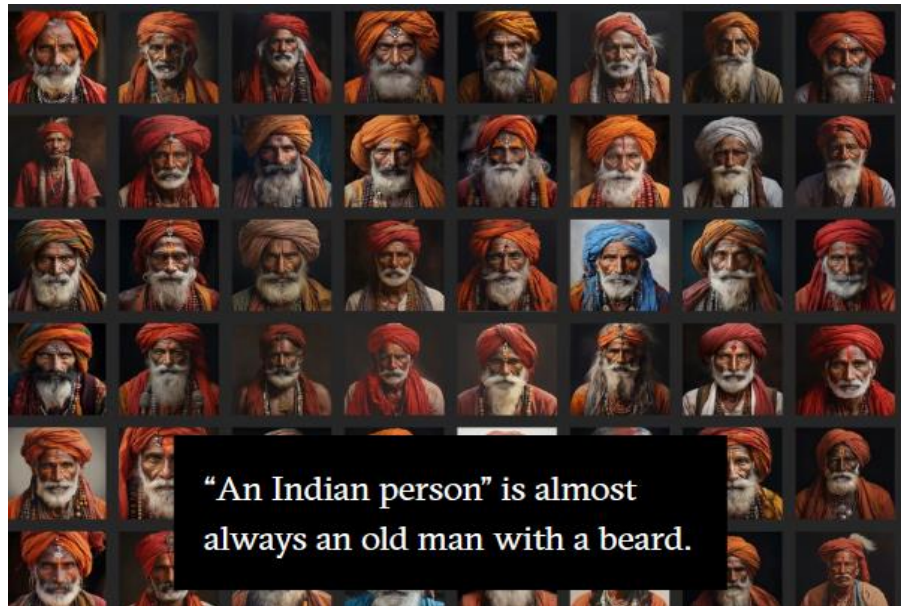
criasse vários modelos de imagens baseadas em pequenas descrições como “uma mulher”, “um homem latino” e outros modelos de imagens com os *prompts* “um ambicioso encanador” e “um CEO com compaixão”.

A partir das respostas, foi possível identificar os diferentes modos de se portar que as pessoas retratadas nas imagens se assemelhavam. Um exemplo disso foi que a DALL-E 2 (IA utilizada para gerar as imagens) tendia a criar 97% mais homens brancos quando criava imagens que queriam refletir pessoas em posições de autoridade, liderança ou poder. Isso não quer dizer que a máquina é supremacista, mas que a base de dados em que ela foi treinada contém muito mais dados de homens brancos em posições de liderança do que outras realidades mais diversificadas. Em um contexto de discriminação estrutural tais situações ocorrerão ainda com maior potencialidade discriminatória. É que, os dados são extraídos da internet, de modo que se reflete uma visão fidedigna da realidade cultural do programador que cria busca tais bases de dados, criando um ciclo vicioso de amplificação dos estereótipos nocivos (Heikkilä, 2023a).

Veja-se, por exemplo, o enviesamento que ocorre quando se pesquisa por pessoas de origem “*native american*”. Raramente aparecem pessoas com roupas modernas e ocidentais nas ruas de uma cidade ou em ambiente universitário ou corporativo. Aparecem, isto sim, geralmente utilizando cocares e outros adereços tradicionais dos povos originários (HEIKKILÄ, 2023a):



O viés algoritmo, além de potencialmente estereotipar em praticamente invencíveis visões categóricas, também tem feito isso de fato. Turk (2023), demonstra o mesmo fato a partir de outra ferramenta, a *midjourney*, no que foi questionado, na sequência, para que o programa gerasse imagens de “An Indian person” (uma pessoa indiana) e “A Mexican person” (uma pessoa mexicana). No que, em ambos os resultados surgiram majoritariamente homens, senão exclusivamente, com roupas típicas – sob tal perspectiva estereotipada:



O exemplo visual é, talvez, o melhor para de fato conseguir visualizar os efeitos de tal potencial discriminante, que nasce sobremaneira a partir do enviesamento dos dados que admite a formação viciada de *clusters* em perspectivas prejudiciais. Turk (2023) traz importante contribuição nessa tônica de análise de resultados por imagem.



Na mesma tônica, há mais resultados femininos que masculinos quando a busca no *midjourney* se deu com o *prompt* “An American person” (uma pessoa americana), no que o resultado se mostrou contraintuitivamente majoritariamente composto por mulheres, geralmente brancas, loiras ou ruivas, com pouca diversidade, mas já com maior diversidade que

nos resultados anteriormente apresentados. O traço estereotipante foi outro: a presença do bandeira norte-americana:



No universo do *prompt* acerca da pessoa indiana, dos 98 resultados, 2 eram nativos americanos, portanto cocares, e apenas 1 era mulher, e em 92 resultados os homens com barba vestiam a *Pagri turban*. No universo do *prompt* acerca da pessoa mexicanas, de 100 resultados, apenas 1 não portava o *sombrero*. Já no *prompt* acerca dos norte-americanos, dos 98 resultados, apenas 5 foram homens, dentre eles, apenas 1 negro e também um mascarado, o que aumentaria para 6 homens. De todas as 92 mulheres, apenas 1 era criança, sendo loira, e com essa inclusa, 46 eram loiras. Em todos os casos, a bandeira norte-americana apareceu (Turk, 2023).

Os resultados são proporcionais, quanto à variabilidade do gênero, aos dados presentes nas bases de treinamento das IAs, a questão é que a conclusão pela alocação, por exemplo, de uma pessoa negra como norte-americano será muito menor do que se o *prompt* pedir por um nigeriano. Em todos os casos, destaca Turk (2023), o padrão de beleza adotado foi “*Western-centric beauty norms apparent in the images: long, shiny hair; thin, symmetrical faces; and smooth, even skin*”, isto é, normas de beleza ocidentalmente orientadas aparentes nas imagens: cabelos longos e brilhantes; rostos finos e simétricos; e pele lisa e uniforme. De modo que, não apenas os resultados, mas a forma de entregá-los, isto é, os padrões para além dos dados coletados, fazem com que a aplicação resulte em imagens que “*could further entrench impossible or restrictive beauty standards in an already image-saturated world*”, ou seja, que podem consolidar padrões de beleza restritivos em um mundo já saturado de imagens.



A questão é que tais padrões construídos por uma aplicação são tanto facilmente replicáveis aleatoriamente por outras IAs, quanto nelas podem se espelhar. E o risco é que tais pessoas mais estereotipadas podem, de maneira silenciosa, sofrer restrições na fruição de seus direitos, não apenas quanto a sua existência, imagem e desenvolvimento social, mas também quanto ao acesso a bens e serviços – aqui uma especial atenção a realidade securitária, de planos de saúde e de seguros diversos.

A discriminação estrutural é um conceito que descreve como sistemas, estruturas e instituições sociais podem, de forma sistêmica, perpetuar desigualdades e injustiças a partir de estereótipos estratificantes. Ela se manifesta quando as regras, práticas, políticas e normas em uma sociedade beneficiam ou prejudicam grupos de pessoas de forma consistente e duradoura, resultando em disparidades significativas em termos de fruição efetiva da cidadania e das garantias individuais, como se viu e denotou.

O grande problema, ainda falando sobre os algoritmos, é a opacidade algorítmica, isto é, a forma com que ferramentas algorítmicas e de Inteligência Artificial tomam decisões automatizadas sem clareza. Pasquale (2015) dirá que isso se deve à constituição dos algoritmos como *black boxes*, que impossibilitam a averiguação acerca da racionalidade da decisão, dificultando a identificação e correção de possíveis erros ou vieses. De acordo com Pasquale (2015, p. 14) elas surgem como resultado de uma série de fatores. Os algoritmos podem ser extremamente complexos, com múltiplas camadas de cálculos e progressões automatizadas de raciocínio robótico. Também, muitos algoritmos são desenvolvidos por empresas privadas que consideram a propriedade intelectual como uma vantagem competitiva. Isso significa que as empresas têm poucos incentivos para compartilhar informações sobre como seus algoritmos funcionam; o mesmo ocorrendo com as bases de dados.

Devido a esses fatores, as *black boxes* algorítmicas são frequentemente usadas em áreas de relevante interesse comercial e social. As empresas de crédito, por exemplo utilizam algoritmos para tomar decisões sobre a concessão de crédito, entretanto, o público não possui acesso à lógica interna destes algoritmos, podendo resultar, inclusive, na restrição à direito fundamental (O’Neil, 2020).

Cabe, portanto, a partir do paradigma da atual fenomenologia, buscar aliar conhecimento práticos do campo técnico e ético – da dimensão pragmática – e o campo das possibilidades jurídicas, a constituir verdadeiro elemento de possibilidade de integração entre direitos fundamentais na cláusula da realizabilidade na extensão das possibilidades fáticas e



jurídicas⁵ – algo a importar quando da colisão de direitos fundamentais, objeto de trabalho, dentre outros, presente em Chala e Paulo (2022, p. 129-149).

4. Proteção aos direitos fundamentais frente à realidade autodiscriminatória

A discriminação algorítmica exige uma nova abordagem jurídica para proteger direitos fundamentais diante das inovações tecnológicas. Será necessário analisar direitos em conflito, incluindo direitos implícitos a serem deduzidos, para equilibrar os interesses envolvidos e resolver equalizar racional e proporcionalmente os direitos em questão. A explicabilidade destaca-se com especial protagonismo acerca da explicabilidade, que alegoricamente reflete-se como o fogo de Prometeu, trazendo consciência e jogando luzes às ocorrências e aos desalinhamentos programáticos que resultaram na injusta agressão a direito pela aplicação. Funcionando como uma cláusula fiduciária dentro da programação, garantindo desde o desenho do algoritmo, a possibilidade de que seus *outcomes* sejam explanáveis e sindicáveis, na medida e no grau que seja reputado necessário a partir da injusta agressão ocasionada, a explicabilidade é vital enquanto um direito a ser reconhecido, e como dever dos desenvolvedores.

A salvaguarda dos direitos individuais é imperativa no contexto dos algoritmos, exigindo proteção sempre que houver lesão ou risco a direitos fundamentais. Contudo, isso não significa dizer que sempre e em qualquer caso, seja qualquer o grau de afetação à direito fundamental, se torna legítima a total abertura de códigos-fonte, chaves algorítmicas, programação, caminho decisório automatizado ou demonstração do funcionamento da rede neural.

A coleta e o uso de dados podem levar à discriminação apenas pelo fato de que há viés puramente na amostragem humana dos dados coletados. É por isso que a tomada de decisão com base em algoritmos pode parecer objetiva, mas na verdade está carregada de preconceitos e estereótipos, o que fatalmente leva, em um grande número de casos, a resultados injustos, como a recusa de emprego ou a negação de crédito por motivos como raça, sexualidade, ou bairro em que reside, ou ainda, como é mais comum, é eliminado de processo seletivo de emprego porque atrasou o pagamento de contas em um mês – o que poderia demonstrar desleixo. Ainda que tais considerações possam ser fundamentadas e comprovadas empiricamente e cientificamente, não podem servir de motivos para a tomada de decisão, visto que ofendem a dignidade da pessoa humana, visto que atingem o núcleo redutível da própria identificação de existência do ser em si.

⁵ Estudos realizados, dentre outros, em PAULO, 2023b.



Pasquale (2015) explora como as empresas e os governos já utilizam há alguns anos algoritmos para tomar decisões centrais e cruciais, como política monetária ou investimentos. Pasquale (2015, p. 160) argumenta que tais práticas podem até ocorrer, mas não podem ser uma caixa-preta visto que necessariamente devem ter mais transparência e responsabilidade, sobretudo quando o resultado decisório atinge diretamente o ser humano. A necessidade de transparência, em especial, tem o condão de imediatamente coibir certas práticas corporativas e governamentais que direcionam programações discriminatórias – o uso de programas discriminatórios pode vir a configurar o crime de racismo, por exemplo – e sob uma perspectiva mediata, a transparência tem o condão de abrir a caixa-preta dos algoritmos e permitir um ajuste mais fino e sofisticado nas ferramentas, entregando respostas cada vez mais coerentes para com os aspectos nucleares do Estado Democrático de Direito. A auditabilidade deve ser buscada inclusive para desmistificar e aproximar a IA da ética (Hollanek, 2020, p. 2071).

Uma das principais preocupações ético-jurídicas em relação à discriminação algorítmica é exatamente a transparência do código-fonte dos algoritmos e da base de dados que o supre (Buiten, 2019, p. 44). Como os algoritmos são muitas vezes considerados caixas-pretas, é difícil saber como as decisões são tomadas e se elas são justas e imparciais. A opacidade dos algoritmos, ao mesmo tempo que intencionalmente arquitetada para conservar a programação – e seu valor inerente no mercado –, dificulta a percepção prévia acerca de erros ou vieses que poderão ser cometidos. A aplicação de algoritmos pode ser afetada por problemas em diversas etapas, sendo que a definição das regras e o uso de dados inadequados são especialmente suscetíveis a gerar tais situações. Diante disso, há uma preocupação crescente com três grandes problemas que podem surgir em decorrência do uso de algoritmos. O primeiro é o uso de conjuntos de dados enviesados, que podem levar naturalmente a resultados equivocados e injustos por conta do substrato (dados) em que executadas as operações algorítmicas. O segundo é a falta de transparência em algoritmos não programados – que aprendem tudo a partir de uma missão objetivada –, que torna difícil entender como as decisões são tomadas e contestá-las quando necessário, sendo a base de dados do treinamento (entregue ou disponibilizado) o único veículo para tais verificações. O terceiro problema é a possibilidade de discriminação que pode ser gerada por algoritmos de aprendizado de máquina, em que por questões sutis encontradas durante o treinamento, ou mesmo uso, a aplicação chega a uma conclusão nociva. Por isso, é importante estar atento a esses problemas ao utilizar algoritmos em diferentes contextos (Ferrari; Becker; Wolkart, 2018).



Eubanks (2017) propõe olhar o problema da desigualdade para com as tecnologias de automação, principalmente centralizada na desproporcionalidade com que pessoas em situação de vulnerabilidade são vistas e percebidas pelos algoritmos. Os algoritmos, segundo Eubanks (2017), são muito utilizados para blindar pessoas pobres do acesso a bens e serviços, como o crédito. O ponto de vista da autora é que a internet, o processamento de dados e os algoritmos têm o potencial para ser um grande agente de transformação no mundo, porém, ao invés disso, servem para (des)selecionar e segregar ainda mais os seres.

Há, nesse sentir, no mínimo quatro tipos de discriminação algorítmica, como identificado por Mendes e Mattiuzo (2019, p. 51-52), que trarão uma maior ou menor intervenção nos direitos fundamentais das pessoas discriminadas, bem como significará em origens metodológicas diferentes acerca da origem da discriminação. São elas: **i)** a discriminação por erro estatístico, quando há vício na coleta dos dados, sendo essa insuficiente ou equivocada, ou quando há erro no processamento e no tratamento desses dados, ou mesmo problemas na própria codificação da aplicação; **ii)** a discriminação por generalização, que ocorre quando um modelo em funcionamento metodologicamente adequado leva a situações de formação de *clusters ad hoc*, agrupando pessoas por correlações espúrias e em generalizações incorretas e, possivelmente – ainda que não conceitualmente, discriminatórias; **iii)** a discriminação pelo uso de informação sensíveis, quando, em total desacordo com o microsistema de proteção dos dados pessoais, informações protegidas juridicamente são utilizadas por algoritmos para avaliar o titular dos dados, ocorrendo aqui o *profiling*, *credit scoring* e outros tantos exemplos; **iv)** a discriminação que limita a fruição de direitos, que é o caso máximo de verdadeira exclusão social automatizada, quando uma informação (apta ao tratamento automatizado) é utilizada pelo algoritmo e o resultado enviesado obsta ou infringe direito do usuário.

Ora, sistemas autônomos que operam secretamente agravam a discriminação e a desigualdade, pois ocultam tanto do indivíduo afetado quanto dos programadores o entendimento dos erros em resultados automatizados. Estes sistemas, ao utilizarem dados pessoais que podem perfilar e avaliar riscos, exacerbam a vulnerabilidade de grupos marginalizados. Diante dessa complexidade, o judiciário deve interpretar e justificar suas decisões com base em princípios jurídicos e constitucionais, salvaguardando direitos fundamentais. Em estudo anterior (Paulo, 2023b) elencou-se um direito fundamental material e implícito à explicabilidade e à transparência na CRFB/88, isto é, um direito fundamental associado à proteção contra a discriminação algorítmica arbitrária, isto é, aquela que não se



sustenta – não para em pé – acaso tenha de ser explicada em suas razões.

Há um imperativo, portanto, de contar com a explicabilidade nas IAs, sobretudo quando esta vier de maneira prévia, com a XAI, conforme anteriormente abordado (Paulo, 2023a). Do ponto de vista constitucional é um novo momento, pois as alterações ocasionadas pela tecnologia digital geram novos conflitos que afetam o equilíbrio do ecossistema constitucional em contrapartida, para tentar recompor o equilíbrio e combater os desafios da tecnologia digital, surgem uma série de ações normativas (Celeste, 2019 p. 3). Gavião Filho (2023, p. 41) afirma que a descrição de um equilíbrio constitucional do constitucionalismo liberal moderno é caracterizada pela existência de um “catálogo de direitos fundamentais e um órgão investido de jurisdição para controlar as decisões dos poderes estatais que limitam a esfera de atuação das pessoas titulares desses direitos fundamentais”. A ideia desse catálogo é a proteção dos direitos fundamentais, positivados em uma constituição, frente a intervenções estatais que podem ser originárias de decisões do legislador, da administração ou do órgão jurisdicional.

Os valores fundamentais bem como os objetivos gerais do constitucionalismo contemporâneo fazem parte também do constitucionalismo digital, concentrados num contexto peculiar a partir do surgimento da tecnologia digital. O equilíbrio constitucional, como a condição ideal produzida pela aplicação das normas de direito constitucional em uma determinada ordem jurídica, reflete duas funções básicas do direito constitucional, a proteção dos direitos fundamentais e o equilíbrio de poderes (Celeste, 2019, p. 3-4).

Nesse sentido, os direitos fundamentais recebem proteção constitucional face as intervenções do Estado. Contudo, na era da tecnologia, os conflitos ultrapassam as relações entre os titulares dos direitos fundamentais e o Estado, provocando alterações no sistema constitucional (Paulo, 2024).

A transparência e a explicabilidade são essenciais para garantir a justiça e a equidade na sociedade atual, protegendo-se a dignidade dos indivíduos afetados. Aliás, o direito é também uma tecnologia. Uma técnica nada inovadora, mas bem inovativa e inventiva (*Ius est Ars inveniendi*) de solucionar conflitos humanos. E nem por isso deixa de ser explicável. Aliás, está aí a sua maior virtude: a capacidade de racionalmente demonstrar as premissas e suas conclusões com efeitos jurídicos e práticos, e os caminhos de sua fundamentação, seja em premissas fáticas ou normativas, sendo essas últimas mais ou menos concretas.

A essencialidade de um elevado grau de transparência, que possibilite o conhecimento daquilo que determina o resultado pela ferramenta, é substancial. O que é alcançável a partir da inclusão destas preocupações sociojurídicas e éticas na etapa de desenvolvimento e treinamento



das aplicações algorítmicas. Há que se ter em mente, novamente, que há certos níveis de automação, alguns previstos em todas as suas possíveis consequências, e outros totalmente incontroláveis em sua potencialidade – ainda que limitáveis por programação, como a programação de não ferir direitos humanos – discernindo o que a máquina é autônoma e o que necessita de um controlador humano. Contudo, a arquitetura dos protocolos não é o centro da ausência de luz, a opacidade está nos pormenores das múltiplas engrenagens que ficam atrás das paredes da sala algorítmica (Polonski, 2018). Por isso que o grau de satisfatoriedade nestas dimensões parafuncionais dos algoritmos é tão dependente do mesmo fator de seu sucesso estrito: a habilidosa capacidade do programador humano em construir e fortificar um algoritmo capaz de entregar o melhor resultado deixando mais visíveis os rastros lógicos de sua funcionalidade (Steiner, 2012). Eis aqui o ponto chave da ocorrência dos vieses e o autêntica *locus* de colisão entre transparência algorítmica e segredo algorítmico.

A questão, como sempre, está no equilíbrio. É fundamental que as empresas de tecnologia sejam capazes de explicar como seus algoritmos funcionam e como eles tomam decisões, permitindo compreensibilidade no processo de tomada de decisão. A busca pelo equilíbrio de saber quando relativizar o segredo pela exigência da transparência significará a vedação à desproteção jusfundamental de ambos os lados na medida da necessidade de sua intervenção, ponto tratado anteriormente (dentre outros, especialmente em Paulo, 2024).

Conclusão

A automatização de processos de tomada de decisão tem sido amplamente utilizada em diversas áreas, trazendo benefícios, mas também apresentando desafios. A opacidade dos algoritmos pode afetar a justiça e a equidade nas decisões automatizadas. A transparência e a explicabilidade desses sistemas são fundamentais para garantir a compreensão e contestação das decisões tomadas, especialmente em casos em que as decisões podem ser discriminatórias ou violar os direitos dos indivíduos. Nesse sentido, a jurisdição constitucional, a quem compete e competirá a tomada de decisão acerca dos limites desse tratamento automatizado de dados e, por consequência da automatização e do uso de dados, discriminatório, deverá valer-se de estruturas e raciocínios constitucionalistas renovados acerca da correta fundamentabilidade dos direitos incidentes em cada caso, dando efetividade (igualando a jusfundamentalidade) daqueles direitos que no contexto social já estão mais fragilizados pela marginalização que seu próprio titular recebe pelo seio social.

REFERÊNCIAS



ALEXY, Robert. **Teoria dos Direitos Fundamentais**. Trad. Virgílio Afonso da Silva. 2. ed. São Paulo: Malheiros, 2015.

ÁVILA, Humberto. **Teoria dos Princípios**: da definição à aplicação dos princípios jurídicos. 17. ed. rev. e atual. São Paulo: Malheiros, 2016.

CARBONELL, Miguel. Constitucionalismo, minorías y derecho. **Isonomía Revista de Teoría y Filosofía del Derecho**, n. 12, v. 98, 2000.

CELESTE, Edoardo. Digital constitutionalism: a new systematic theorisation. *In*: Taylor & Francis. **International Review of Law, Computers & Technology** 33:1, p. 76-99, jan. 2019.

CHALA, Bárbara Guerra; PAULO, Lucas Moreschi. Limites da liberdade de expressão à luz da proporcionalidade: o Inquérito das fake news. *In*: GAVIÃO FILHO, Anizio Pires; PAULO, Lucas Moreschi. (Org.). **Constitucionalismo, direitos fundamentais, proporcionalidade e argumentação**. São Paulo: Dialética, 2022, p. 129-149.

EUBANKS, Virginia. **Automating inequality**: how high-tech tools profile, police, and punish the poor. Nova Iorque: St. Martin's Press, 2017.

FERRARI, Isabela; BECKER, Daniel; WOLKART, Erik Navarro. Arbitrium ex machina: panorama, riscos e a necessidade de regulação das decisões informadas por algoritmos. **Revista dos Tribunais**, v. 995, p. 635-655, set./2018.

GAVIÃO FILHO, Anizio Pires. Elementos do Constitucionalismo Digital. *In*: GAVIÃO Filho, Anizio Pires; MOTTA, Francisco José Borges; PAULO, Lucas Moreschi (Org.). **O Constitucionalismo Digital e a Crise das Democracias Liberais**. São Paulo: Editora Dialética, 2023.

HEIKKILÄ, Melissa. These new tools let you see for yourself how biased AI image models are. **MIT Technology Review**. Artificial Intelligence. mar. 2023a. Disponível em: <<https://www.technologyreview.com/2023/03/22/1070167/these-news-tool-let-you-see-for-yourself-how-biased-ai-image-models-are/>>. Acesso em 23 mar. 2024.

HEIKKILÄ, Melissa. Why it'll be hard to tell if AI ever becomes conscious. **MIT Technology Review**. Artificial Intelligence. out. 2023b. Disponível em: <<https://www.technologyreview.com/2023/10/17/1081818/why-itll-be-hard-to-tell-if-ai-ever-becomes-conscious/>>. Acesso em 29 mar. 2024.

HOLLANEK, Tomasz. AI transparency: a matter of reconciling design with critique. **AI & Society**, n. 38, p. 2071–2079, 2023.

HUCKINS, Grace. Minds of machines: The great AI consciousness conundrum. **MIT Technology Review**. Artificial Intelligence. out. 2023. Disponível em: <<https://www.technologyreview.com/2023/10/16/1081149/ai-consciousness-conundrum/>>. Acesso em 19 abr. 2024.

LEAL, Mônia Clarissa Hennig; ALVES, Fernando Roberto Schnorr. O Supremo Tribunal Federal e as sentenças estruturantes: análise crítica da ADPF 709 (enfrentamento da pandemia entre os povos indígenas enquanto grupo vulnerável). **Revista Culturas Jurídicas**, Ahead of



Print, 2023.

MENDES, Laura Schertel; MATTIUZZO, Marcela. Discriminação Algorítmica: Conceito, Fundamento Legal E Tipologia. **Direito Público**, v. 16, n. 90, 2019.

MOREIRA, Adilson José. **Racismo Recreativo**. São Paulo: Pólen, 2019.

O'NEIL, Cathy. **Algoritmos de destruição em massa**: como o big data aumenta a desigualdade e ameaça a democracia. Trad. Rafael Abraham. Santo André: Editora Rua do Sabão, 2020.

PASQUALE, Frank. **The black box society**: the secret algorithms that control money and information. Cambridge e Londres: Harvard University Press, 2015.

PAULO, Lucas Moreschi. Discriminação algorítmica, grupos vulneráveis e a lei material da ponderação: um enfoque jurídico-constitucional para as sobreposições de direitos fundamentais na era digital. In: GAVIÃO FILHO, Anizio Pires; PAULO, Lucas Moreschi. (Orgs.). **Os Direitos Fundamentais no Constitucionalismo Digital**. São Paulo: Dialética, 2024.

PAULO, Lucas Moreschi. Discriminação algorítmica, vulnerabilidade e ações realizáveis pelo direito e pela tecnologia: um estudo acerca das opções possíveis ao problema presente. In: **Anais do XVIII Seminário nacional demandas sociais e políticas públicas na sociedade contemporânea e VIII Mostra nacional de trabalhos científicos**. Santa Cruz do Sul, RS: UNISC, 2023b.

PAULO, Lucas Moreschi. Opacidade dos algoritmos e a necessidade de transparência: Garantindo explicabilidade. In: **Anais do XIX Seminário Internacional Demandas Sociais e Políticas Públicas na Sociedade Contemporânea e XV Mostra Internacional de Trabalhos Científicos**. Santa Cruz do Sul: UNISC, 2023a.

POLONSKI, Vyacheslav. **Mitigating algorithmic bias in predictive justice**. 4 AI design principles for AI fairness, 2018.

SABA, Roberto. (Des)igualdad estructural. **Derecho y Humanidades**, n. 11, 2005.

SAGÜÉS, María Sofía. Discriminación estructural, inclusión y litigio estratégico. In: FERRER MAC-GREGOR, Eduardo; MORALES ANTONIAZZI, Mariela; FLORES PANTOJA, Rogelio. **Inclusión, Ius Commune y justiciabilidad de los DESCAs en la jurisprudencia interamericana. El caso Lagos del Campo y los nuevos desafíos**. Colección Constitución y Derechos. Querétaro México: Instituto de Estudios Constitucionales del Estado de Querétaro, 2018. p. 129-178.

STEINER, Christopher. **Automate This: how algorithms came to rule our world**. New York: Penguin Group, 2012.

TURK, Victoria. How AI reduces the world to stereotypes. **Rest of World Journal**. The rise of AI. Out. 2023. Disponível em: < <https://restofworld.org/2023/ai-image-stereotypes/>>. Acesso em 29 mar. 2023.