



ORIGINAL

#### Editores

Cláudio José Franzolin, Pedro Pulzatto Peruzzo

#### Conflito de interesses

Os autores declaram não haver conflito de interesses.

#### Recebido

4 jun. 2025

#### Aprovado

16 out. 2025

## REVISTA DE DIREITOS HUMANOS E DESENVOLVIMENTO SOCIAL

# O dever de *accountability* algorítmica sob a Convenção Americana sobre Direitos Humanos

## *Algorithmic accountability duty under the American Convention on Human Rights*

Ana Leticia Szkudlarek<sup>1,2</sup> , Melina Girardi Fachin<sup>1,3,4</sup> 

<sup>1</sup> Universidade Federal do Paraná, Núcleo de Estudos em Sistemas de Direitos Humanos. Curitiba, PR, Brasil.

<sup>2</sup> Tribunal de Justiça do Estado do Paraná. Curitiba, PR, Brasil. Correspondência para: A. L. Szkudlarek. E-mail: <analeszk@gmail.com>.

<sup>3</sup> Universidade Federal do Paraná, Faculdade de Direito. Curitiba, PR, Brasil.

<sup>4</sup> Universidade Federal do Paraná, Pró-Reitoria de Pós-Graduação, Curso de Pós-Graduação *Stricto Sensu* em Direito. Curitiba, PR, Brasil.

**Como citar este artigo:** Szkudlarek, A. L.; Fachin, M. G. O dever de *accountability* algorítmica sob a Convenção Americana sobre Direitos Humanos. *Revista de Direitos Humanos e Desenvolvimento Social*, v. 6, e2516093, 2025. Doi: <https://doi.org/10.24220/2675-9160v6a2025e16093>

### Resumo

O presente artigo busca demonstrar que a Convenção Americana sobre Direitos Humanos impõe ao Brasil o dever de regular, respeitar e supervisionar a *accountability* algorítmica. Para isso, com base na metodologia hipotético-dedutiva, o trabalho explica os aspectos fundamentais dos algoritmos, expõe os seus riscos à democracia e aos direitos humanos e destaca a natureza política desse campo, evidenciando a urgente necessidade de normatização da *accountability* algorítmica no âmbito interno. A partir da perspectiva multinível dos direitos humanos, confirma-se a hipótese de que a Convenção Americana sobre Direitos Humanos exige do Estado brasileiro a regulação, o respeito e a supervisão da *accountability* algorítmica, sob pena de responsabilização internacional, inclusive por omissão.

**Palavras-chave:** *Accountability* algorítmica. Convenção Americana sobre Direitos Humanos. Direitos humanos.

### Abstract

*This article seeks to demonstrate that the American Convention on Human Rights imposes on Brazil the duty to regulate, respect and supervise algorithmic accountability. To this end, based on the hypothetical-deductive methodology, it explains the fundamental aspects of algorithms, exposes their risks to democracy and human rights and highlights the political nature of this field, underscoring the urgent need for domestic regulation of algorithmic accountability. From a multilevel human rights perspective, the hypothesis is confirmed that the American Convention on Human Rights requires the Brazilian State to regulate, respect and supervise algorithmic accountability, under penalty of international liability, including for omission.*

**Keywords:** *Algorithmic accountability. American Convention on Human Rights. Human rights.*



## Introdução

Sentada na calçada da Avenida Nossa Senhora de Copacabana, uma mulher, sem documentos, é surpreendida pela Polícia Militar do Rio de Janeiro e detida. Após familiares levarem sua identidade à 12ª Delegacia Policial, o porta-voz da Secretaria Estadual de Polícia Militar, coronel Mauro Fliess, assume o erro: o sistema de reconhecimento facial, recentemente instalado em câmeras pela cidade, equivocadamente a identificou como uma pessoa supostamente foragida, condenada por homicídio. Entretanto, mais uma falha: a mulher procurada, e pretensamente identificada pelas câmeras de vigilância, está presa desde 2015. Indagado, Fliess destacou que o projeto de reconhecimento facial ainda está em fase de testes. O sistema funciona com base em uma tecnologia algorítmica, que compara características faciais com imagens armazenadas em um banco de dados (Ribeiro, 2023).

Apesar de ser esse o desejo de muitos, não há um momento ao qual retornar: a tomada de decisão algorítmica e a Inteligência Artificial (IA) são uma realidade. A falsa percepção de que esse campo é completamente técnico – quando é fundamentalmente político (Instituto de Tecnologia & Sociedade do Rio, 2024) – tem distanciado as pessoas, especialmente aquelas alheias à área da tecnologia, de tentar compreendê-lo.

Eventual distanciamento social do debate, no entanto, não reflete a intrínseca relação entre os algoritmos e as diferentes dimensões da existência cotidiana humana (Parra; Abdo, 2016): na socialização, com as redes sociais; no transporte e geolocalização, com os mapas via satélite; na segurança, com o reconhecimento facial e biométrico; na culinária, com as entregas por aplicativo; nos serviços bancários, com as transferências virtuais e análises de crédito; na gestão de casas, drones e automóveis inteligentes. Alimentados pelos rastros digitais das ações humanas (Bezerra; Costa, 2022), os algoritmos são utilizados para aprovar ou negar vistos a imigrantes, na contagem de votos em eleições, nos processos de recrutamento de Recursos Humanos (RH) (Köchling; Wehner, 2020), para decidir quais contribuintes passarão por auditoria em seus impostos de renda e mesmo nos sistemas de justiça (Kroll *et al.*, 2016).

As transformações advindas do mundo digital, com o desenvolvimento de sistemas algorítmicos cada vez mais complexos, criam possibilidades, mas também “novas formas de sujeição social e servidão maquínica” (Parra; Abdo, 2016, p. 138). As grandes empresas de tecnologia digital (*as Big Techs*), majoritariamente provenientes do norte-global (Silva, 2020), guiam e moldam comportamentos, impactando profundamente a cultura pública ao redor do mundo.

Nessa conjuntura, cabe analisar e estabelecer limites ao uso dos algoritmos, considerando a premissa de um constitucionalismo multinível. Defende-se “a importância da interface e dos diálogos dos diferentes planos protetivos para a realização dos direitos humanos – o que demanda a relação entre os constitucionalismos e destes com o direito internacional dos direitos humanos” (Fachin, 2021, p. 54).

Nesse contexto, o presente trabalho busca refletir, pela ótica do constitucionalismo multinível – especificamente do Sistema Interamericano de Direitos Humanos (SIDH) –, qual é o papel dos Estados em relação à tomada de decisão algorítmica.

A partir da metodologia hipotético-dedutiva, objetiva-se introduzir, primeiramente, uma compreensão elementar dos algoritmos. Em seguida, pretende-se escancarar as potenciais ameaças dos sistemas de decisões automatizadas às democracias e aos direitos humanos. Em sequência, evidencia-se a urgência de se estabelecer, com suficiente densidade, a *accountability* algorítmica no âmbito do ordenamento jurídico interno. Ao final, a partir da perspectiva multinível

dos direitos humanos, o trabalho busca responder se é possível extrair, da Convenção Americana sobre Direitos Humanos (CADH), o dever dos Estados Partes de adotar disposições internas sobre *accountability* algorítmica (em observância ao artigo 2º da CADH), bem como de respeitá-la (artigo 1º da CADH) e supervisioná-la. Todavia, a viabilidade de um debate sobre o assunto pressupõe compreender o que são os algoritmos.

## Aspectos elementares dos algoritmos

O algoritmo é, em suma, uma sequência de instruções – raciocínios ou operações (Branco; Teffé, 2022) – que diz a um sistema de computação o que fazer e como executar determinada tarefa (Ferrari, 2018) – ou seja, que visa alcançar um objetivo (Branco; Teffé, 2022). Podem ser definidos, ainda, como processos ou conjuntos de regras a serem seguidos em cálculos ou outras operações de resolução de problemas, especialmente por um computador (Köchling; Wehner, 2020). Em acepção complementar, são descritos como “instrumentos usados para converter dados em evidências quanto a um determinado resultado/cenário, que, subsequentemente, é usado para desencadear e motivar uma ação que pode ter consequências que ensejem uma determinada apreciação ética” (Wimmer; Doneda, 2021, p. 337).

São os algoritmos a base para diversas ferramentas de decisão de inteligência artificial. Destaque-se que a IA abrange modelos, métodos e prescrições usadas para simular a inteligência humana, muitas vezes quando se trata de coletar, processar e agir com base em dados (Köchling; Wehner, 2020).

Quanto às espécies ou classificações algorítmicas, tem-se os algoritmos programados – ou seja, aqueles que seguem operações definidas pelo programador – e os não programados, isto é, aqueles que possuem capacidade de aprendizagem e criam outros algoritmos. De modo bastante simples, neste último caso, o computador escreve a própria programação (Ferrari, 2018). Os algoritmos não programados são também chamados de técnicas de aprendizagem de máquina ou *machine learning* (Ferrari, 2018).

Entre os algoritmos não programados, encontram-se os supervisionados, os não supervisionados e os de reforço (Ferrari, 2018). No primeiro caso, o sistema é alimentado com dados lapidados e previamente escolhidos por seres humanos. Carregam-se, no computador, os dados já rotulados e as saídas desejadas. O modelo, então, ajusta as variáveis para mapear as entradas (*inputs*) e saídas (*outputs*). Os *learners* supervisionados visam fazer previsões a partir da aprendizagem de padrões, apreendida por meio dos dados rotulados fornecidos por humanos (Köchling; Wehner, 2020).

Quando se trata, porém, de algoritmos não supervisionados – e é aqui que residem as maiores problemáticas de uma explicabilidade efetiva da tomada de decisão algorítmica –, os dados que alimentam o sistema não são rotulados, ou seja, não há uma classe de dados pré-definida. Para a análise ou o agrupamento de temas de dados, os algoritmos de *machine learning* não supervisionados capturam os comportamentos estruturais de variáveis nos dados de entrada (Köchling; Wehner, 2020). São muito úteis, portanto, à descoberta de padrões, estando presentes no reconhecimento e identificação facial e de voz, na criação de carros e drones autônomos, bem como no diagnóstico de doenças.

Por fim, os algoritmos de reforço não se baseiam em dados fixos de entrada e saída. Em vez disso, aprendem o comportamento por meio de interações de tentativa e erro com um ambiente dinâmico (Köchling; Wehner, 2020) e enfocam na performance – são bastante utilizados em jogos e redes sociais (Ferrari, 2018).

As tarefas desempenhadas pelos algoritmos não programados – mesmo no caso dos supervisionados – são difíceis de prever, em especial quando se está diante dos algoritmos não supervisionados e de reforço. Segundo Ferrari (2018), em menção a Burrell (2016), a autonomia dessas espécies algorítmicas, a alta dimensionalidade de dados, a complexidade dos códigos e das combinações probabilísticas de suas previsões, a variabilidade da lógica de tomada de decisões, a velocidade de processamento das informações e a multiplicidade de variáveis operacionais – visto que o algoritmo não programado modifica, autonomamente, a sua própria estrutura – trazem, em conjunto, uma inerente opacidade sobre a forma como operam, tornando difícil compreender os seus processos internos. Os próprios desenvolvedores, muitas vezes, não entendem adequadamente como alguns algoritmos funcionam (Branco; Teffé, 2022).

Em outros termos, a mera abertura do sistema algorítmico – conforme proposto pela organização *OpenAI* – não é medida capaz de enfrentar a centralidade da produção tecnológica e os efeitos deletérios da integração dos algoritmos nos domínios mais sensíveis da vida pública e privada (Parra; Abdo, 2016). Por isso, Ferrari acredita que as alternativas mais promissoras, ao enfrentamento dos efeitos negativos dos algoritmos, devem focar em explicabilidade, não em acessibilidade ao código-fonte (Ferrari, 2018) – a qual envolve a polêmica de se permitir (ou não) a abertura da “caixa-preta” (em inglês, “*black box*”) dos algoritmos. De qualquer maneira, não há caminho isento de intensos debates.

Embora a discussão sobre como conferir explicabilidade aos algoritmos ainda permaneça uma incógnita, já há algumas alternativas interessantes, no sentido de se garantir um mínimo controle, transparência e supervisão da tomada de decisão automatizada.

Antes disso, contudo, devem ficar nítidos os perigos que os algoritmos representam para a sociedade e para a democracia, especialmente à brasileira – a justificar o dever de *accountability* que se propõe.

## **Ameaças algorítmicas**

Já é globalmente reconhecido, inclusive por autoridades políticas (Lopes; Paola, 2024) e pelos próprios laboratórios de IA (Ordonez; Dunn; Noll, 2023), que há profundos riscos da tomada de decisão algorítmica para a sociedade. Assim como as tecnologias antecedentes ao digital “já possuíam caráter político e racializado” (Bezerra; Costa, 2022, p. 5), as avaliações ou previsões algorítmicas acabam por discriminar, sistemática e injustamente, certos indivíduos ou grupos já vulnerabilizados e marginalizados (Köchling; Wehner, 2020).

Cathy O’Neil, cientista de dados, nomeia os modelos algorítmicos nocivos – ou seja, aqueles que reforçam padrões discriminatórios e a desigualdade social –, geralmente opacos e invisíveis, de “Armas de Destruição Matemática” (ADMS). Vendidas como neutras, imparciais e objetivas, as ADMS se autoperpetuam, em larga escala, e são altamente destrutivas (O’Neil, 2020).

Devido à incorporação consciente ou inconsciente de vieses humanos (Ferrari, 2023) – aqui compreendidos como “impressões equivocadas, tendências errôneas e preconceitos” (Ferrari, 2023, p. 117), proporcionadas por premissas duvidosas e simplificações da realidade, de conotação necessariamente negativa –, os dados que alimentam os algoritmos refletem os preconceitos socioculturais e estruturais existentes e, conseqüentemente, podem perpetuá-los em suas decisões. A questão se agrava quando se percebe que os algoritmos podem conferir um caráter de “cientificidade” a decisões enviesadas, normalizando os vieses propagados (Ferrari, 2023). Tal situação é especialmente preocupante em uma sociedade profundamente discriminatória e desigual, como é a brasileira.

O racismo, elemento estrutural da sociedade, também está presente nos dados e, portanto, na tomada de decisão algorítmica, pois “o mundo algoritmizado não existe fora da realidade” (Bezerra; Costa, 2022, p. 4). A diferença é que a desigualdade *datafícada* “esconde a figura do sujeito racista a ser denunciada” (Bezerra; Costa, 2022, p. 4), potencializando as discriminações e dificultando o enfrentamento das assimetrias estruturais existentes.

Com efeito, a análise de mais de 2.512 fotografias e ilustrações, realizada por Carrera e Carvalho (2020), confirmou que as fórmulas algorítmicas reproduzem e fortalecem uma estrutura social racista e machista. Corroborando as ideias de Noble (2018), as pesquisadoras concluíram que mulheres negras são mais representadas sozinhas com seus filhos do que mulheres brancas, e que a busca da palavra-chave “*family*”, em bancos pagos de imagens digitais, resulta em maioria expressiva de famílias brancas. Ao reforçarem a “hiper-ritualização imagética do preterimento afetivo aos negros, sobretudo à mulher negra” (Carrera; Carvalho, 2020, p. 111) e perceberem como “universais” alguns corpos e sujeitos (brancos), os algoritmos se despem de sua carcaça de neutralidade e revelam seus vieses discriminatórios.

Na análise de sistemas de reconhecimento de imagens, usados em carros autônomos, outra pesquisa chegou à assustadora conclusão de que “a acurácia na identificação de pessoas com pele escura poderia ser 5% menor, resultando em potenciais atropelamentos” (Silva, 2020, p. 439). Não há falsos alarmismos, portanto, em afirmar que algoritmos colocam o direito à vida de corpos não-brancos em potencial risco; no entanto, o patrimônio trilionário das *Big Techs*, superior ao Produto Interno Bruto (PIB) de diversos países (Sfredo; Lopes, 2020), cala, com êxito, as tentativas de fiscalização e responsabilização.

Diante da problemática, surge a seguinte questão: como enfrentar os efeitos perversos dos sistemas de decisão automatizada? Para combatê-los, é preciso, primeiramente, entender que a discriminação algorítmica pode advir tanto dos dados empregados para treinar os algoritmos quanto das escolhas associadas à programação desses algoritmos (Ferrari, 2023).

A respeito, Ferrari (2023) explica que pode haver (i) uma seleção equivocada de dados – de forma errônea ou incompleta –, (ii) a absorção, pela máquina, de vieses implícitos da sociedade, presentes nos dados produzidos de maneira difusa e (iii) a absorção de padrões discriminatórios das próprias pessoas decisoras envolvidas com o sistema. Outro problema que pode surgir se dá quando o algoritmo não é capaz de incorporar, na concepção do sistema, novos conhecimentos ou percepções da sociedade, ocasionando a reprodução de preconceitos emergentes (Köchling; Wehner, 2020).

Além disso, não apenas a sub-representação nos dados, mas também a super-representação pode ser prejudicial e gerar ameaças a grupos em maior vulnerabilidade (Ferrari, 2023). Nesse ponto, Ferrari (2023) dá um exemplo interessante: se um conjunto de dados indicar a um algoritmo, desenvolvido para um sistema de justiça, que um determinado grupo de pessoas costuma receber condenações mais severas, é possível que o sistema automatizado passe a reproduzir o padrão encontrado, perpetuando desigualdades existentes.

Por isso, a qualidade dos dados de entrada é fundamental para auxiliar no enfrentamento da discriminação algorítmica (Köchling; Wehner, 2020). Por exemplo, em um processo de recrutamento para uma empresa, se um algoritmo for treinado com dados de históricos de emprego – mesmo sem receber informações sobre gênero ou etnia –, o sistema pode identificar padrões que indicam que determinados grupos têm menos probabilidade de serem selecionados e, conseqüentemente, replicar essa desigualdade ao fazer novas seleções. Essa escolha algorítmica pode levar populações

vulnerabilizadas a desvantagens sistemáticas, mesmo que não seja a intenção da ou do *designer* de dados (Köchling; Wehner, 2020).

Nota-se, portanto, que uma escolha pouco criteriosa dos bancos de dados pode levar à reprodução e ao fortalecimento das injustiças sociais (Branco; Teffé, 2022).

Mas quem está selecionando e treinando esses dados?

De acordo com pesquisadores, os algoritmos são criados, em sua maioria, por programadores brancos do sexo masculino (Instituto de Tecnologia & Sociedade do Rio, 2024; O’Neil, 2020; Bezerra; Costa, 2022). A ausência de pluralidade no setor tecnológico, inequivocamente, reflete-se na cisgeneridade heteronormativa e branca das decisões automatizadas<sup>5</sup>. Exemplo evidente disso envolveu a empresa *Amazon*, que constatou desvantagem extrema às mulheres, em seu algoritmo de contratação<sup>6</sup> (Dastin, 2018; Köchling; Wehner, 2020).

Além da desigualdade de gênero e do racismo, outras opressões são fomentadas pelos sistemas automatizados. A discriminação etária – ou *etarismo* –, por exemplo, está presente em muitas das tomadas de decisões algorítmicas: no *Facebook*, empresas utilizam anúncios direcionados para atrair candidatos a vagas de emprego e, intencionalmente, determinam as características pessoais desejadas, de modo a excluir pessoas mais velhas de sua visualização (Köchling; Wehner, 2020). Já outra pesquisa comprovou a utilização de algoritmos de contratação que consideram a distância entre o local de trabalho e a casa da pessoa como forte preditor de permanência no emprego, o que é nitidamente discriminatório, tendo em vista os processos de gentrificação e distanciamento das populações marginalizadas para as periferias (Köchling; Wehner, 2020).

Restam evidentes, portanto, os perigos da tomada de decisão algorítmica, sem as devidas limitações, aos direitos humanos e à democracia.

Apesar da escassa transparência e explicabilidade dos modelos algorítmicos, sobretudo quanto aos algoritmos de *machine learning*, há alternativas viáveis e que podem mitigar, significativamente, os efeitos deletérios da tomada de decisão automatizada.

No ponto subsequente, será explicado o que é a *accountability* algorítmica, quais são as principais técnicas que podem ser empregadas e o porquê de ser inadiável a sua regulamentação e supervisão.

## **Accountability algorítmica**

Apesar das críticas ao emprego do termo *accountability* – devido à alegada imprecisão terminológica (Mota *et al.*, 2021) –, a ausência de uma expressão, em língua portuguesa, que consiga refletir a abrangência do conceito, levou à adoção do termo, em inglês, neste texto. Destaca-se, ainda, que a opção pelo uso da expressão estrangeira “tem sido uma convenção terminológica não só na literatura das ciências sociais brasileiras, mas também em outras línguas não anglofônicas” (Rached, 2021, p. 190).

Fala-se, aqui, em *accountability* como “prestação de contas” ou “responsabilização” – sem maiores aprofundamentos terminológicos, nesta oportunidade, quanto a eventuais distinções entre esses dois conceitos –, estando intimamente ligada à transparência e a mecanismos ou estruturas de supervisão (Saldanha, 2024) (*oversight*) (Ferrari, 2023). Explica Dawn Oliver que o referido termo

<sup>5</sup> A Resolução 332/2020 do CNJ, em seus arts. 6º e 20, valoriza a composição diversa das equipes que atuam com dados, reconhecendo que isso contribui para a qualidade e integridade das informações (Ferrari, 2023).

<sup>6</sup> Após a constatação, a empresa decidiu encerrar a tomada de decisão algorítmica para essa função.

se relaciona “à exigência de uma pessoa explicar e justificar – com base em critérios de algum tipo – suas decisões ou atos” e reparar, depois, “falhas ou erros” (Regulação Algorítmica, 2020).

Quer-se pensar, a partir da perspectiva anunciada, o papel dos Estados em garantir que as partes envolvidas, na tomada de decisão algorítmica – como as empresas que contratam engenheiros de dados para desenhar modelos matemáticos próprios –, sejam responsáveis pelos seus impactos. Buscam-se, também, efetivos mecanismos de supervisão legal, que assegurem maior transparência e gerenciamento de riscos (Nem [...], 2023), uma vez que as empresas, em geral, não revelam suas práticas (Köchling; Wehner, 2020).

O debate sobre *accountability* algorítmica tem ganhado relevante espaço na cena global: há uma tendência emergente para enfrentar e diminuir os riscos associados a falhas nos sistemas algorítmicos (Wimmer; Doneda, 2021). A apreensão geral com a tomada de decisão algorítmica levou o Chile, por exemplo, a propor uma consulta pública sobre transparência, para estabelecer medidas estruturais a serem seguidas por entidades públicas no assunto. A China, por sua vez, adotou inéditas regulações sobre o uso de tecnologias de recomendação feitas por algoritmos (Lemos, 2021). O governo britânico, de forma ainda mais inovadora, passou a exigir que as organizações do setor público forneçam informações precisas e justificadas sobre as ferramentas algorítmicas que utilizam (Algorithmic [...], 2023). Na Europa, surge um arcabouço normativo interessante sobre o assunto, como o “Ato da IA” (em inglês, *AI Act*), proposto pela União Europeia (European [...], [202-?]) e marcado, principalmente, pela análise de riscos. A *accountability algorítmica* também está na pauta de estudo da *Organization for Economic Co-operation and Development* (OCDE) (Organization for Economic Co-operation and Development, 2023) e, especialmente quanto à equidade algorítmica, é tratada pela Recomendação da Organização das Nações Unidas para a Educação, Ciência e Cultura (UNESCO) sobre a Ética da Inteligência Artificial (Ferrari, 2023).

No Brasil discute-se o Projeto de Lei (PL) nº 2.338 de 2023, que visa à regulamentação da inteligência artificial no país. Com base no texto substitutivo, apresentado em junho de 2024, busca-se: fixar direitos e garantias básicas dos indivíduos potencialmente afetados; definir os usos de IA inaceitáveis (devido aos riscos excessivos); criar obrigações gerais de boa governança; estabelecer medidas de prevenção e combate a discriminações diretas e indiretas, atentando-se à realidade brasileira; definir um arranjo regulatório de fiscalização com autoridade competente e independente; prever a participação social nos processos de governança e de classificação de risco, entre outros objetivos e determinações (Coalizão Direitos na Rede, 2024). Além do referido Projeto de Lei, é interessante mencionar que o Protocolo para Julgamento com Perspectiva de Gênero, elaborado pelo Conselho Nacional de Justiça (CNJ), também traz os perigos da discriminação algorítmica para as mulheres e outros grupos em maior vulnerabilidade social (Conselho Nacional de Justiça, 2021).

Como colocado anteriormente, apesar da opacidade dos modelos algorítmicos, é possível – e necessário – o emprego de algumas ferramentas técnicas para assegurar um nível mínimo de *accountability*, ainda que certos aspectos sejam mantidos em sigilo. É o caso, por exemplo, das técnicas de verificação de *software*<sup>7</sup>, acordos criptográficos<sup>8</sup>, *zero-knowledge proofs*<sup>9</sup> e *fair random choices*<sup>10</sup>.

<sup>7</sup> “[...] garante que, ao operar, o sistema sempre apresentará certas propriedades” (Ferrari, 2018, f. 14).

<sup>8</sup> “[...] asseguram que o programa não foi alterado nem revelado [...] passado certo tempo, dão certeza sobre os critérios utilizados”, podendo-se seguir análise sobre a legitimidade de sua operação pretérita [...]” (Ferrari, 2018, f. 14-15).

<sup>9</sup> “[...] ferramentas criptográficas que permitem que de pronto se prove que a política decisória utilizada apresenta certa propriedade, sem revelar como se sabe disso ou que política decisória é” (Ferrari, 2018, f. 15).

<sup>10</sup> Garantem que “quando o sistema possuir algum nível de aleatoriedade, esta será justa, e não poderá haver intromissão indevida de agentes internos na aleatoriedade do sistema” (Ferrari, 2018, f. 15).

Além das mencionadas técnicas, discutir sobre *accountability* algorítmica envolve refletir, enquanto sociedade, sobre quais modelos matemáticos devem ser priorizados na tomada de determinadas decisões (Ferrari, 2018). No campo do sistema de justiça, por exemplo, é possível que um sistema decisional automatizado não aconselhe adequadamente em litígios (Köchling; Wehner, 2020). Nesse sentido, vale a menção ao questionamento de Isabela Ferrari: “seria legítima a escolha governamental pelo emprego de um algoritmo de *machine learning* do tipo *black box* para escolhas sensíveis” (Ferrari, 2018, f. 14) quando viáveis modelos menos opacos?

Por isso, o controle e a supervisão das decisões algorítmicas devem ser casuisticamente analisados, para o fim de se avaliar: (i) a intensidade; (ii) as formas e (iii) o momento da intervenção humana nesse processo. Algoritmos previsíveis e compreensíveis, como os programados, demandam *accountability* muito mais simples em comparação aos “sistemas do tipo ‘caixa preta forte’” (Wimmer; Doneda, 2021, p. 397).

No mesmo sentido de Wimmer e Doneda (2021), é possível – e, inclusive, necessária – a modulação do grau da intervenção humana, quando se fala da *accountability* algorítmica, em diferentes níveis de intensidade, a depender, entre múltiplos fatores, da transparência, explicabilidade e impacto a direitos e garantias fundamentais de um determinado modelo matemático. Alguns algoritmos podem justificar intervenções humanas mais genéricas ou mesmo indiretas (Wimmer; Doneda, 2021).

Além de o grau de intensidade da *accountability* poder ser distinto, as formas da participação humana, no processo de controle e supervisão algorítmica, também podem variar, conforme a opacidade e o volume de dados.

Soma-se, à análise da intensidade e forma, a apreciação do momento em que a *accountability* será exercida, pois ela pode ocorrer antes ou depois da produção dos efeitos da decisão algorítmica em análise.

Desse modo, a atribuição de tarefas a entes automatizados, em processos decisórios de relevantes impactos, suscita “uma demanda clara pela possibilidade de escrutínio dos diversos elementos deste processo decisório” (Wimmer; Doneda, 2021, p. 376). Surge, então, a reivindicação por uma “espécie de instância na qual estas possíveis ‘falhas’ possam ser verificadas e avaliadas” (Wimmer; Doneda, 2021, p. 376).

Nesse sentido, quais são as alternativas para tentar suprir o “déficit de humanidade” (Wimmer; Doneda, 2021, p. 396) algorítmico?

A implementação de “Avaliações de Impacto Algorítmico” (AIA) pode auxiliar na transparência e mitigação do enviesamento de dados e dos resultados por eles produzidos (Saldanha, 2024), pois permite descrever as características do sistema analisado, bem como identificar riscos e mecanismos para sua mitigação (Mattos Filho, 2023).

No entanto, é possível ir além: os relatórios avaliativos podem evoluir para incluir, em alguns casos, “a previsão de consultas a comitês independentes de especialistas, para apoiar avaliações éticas mais complexas” (Wimmer; Doneda, 2021, p. 396).

Além disso, impor às empresas o dever de publicar, periodicamente, relatórios de transparência, bem como de formular “códigos de conduta” internos, “com base em recomendações emitidas por um órgão regulador”, preferencialmente “composto por representantes de diferentes setores da sociedade” (Instituto de Tecnologia & Sociedade do Rio, 2024, p. 17), são, em conjunto com outras práticas, relevantes alternativas de *accountability*.

Acrescente-se, como mecanismo de *accountability*, que muitos ordenamentos jurídicos, inclusive o brasileiro (artigo 20 da Lei Geral de Proteção de Dados – LGPD), preveem a possibilidade de revisão humana às decisões automatizadas ou, então, “um direito à explicação” (Wimmer; Doneda, 2021, p. 376).

No entanto, os parâmetros normativos existentes, na ordem jurídica nacional, permanecem excessivamente vagos e pouco esquematizados, carecendo de critérios precisos sobre quando se justifica a intervenção humana – ou, minimamente, a revisão, ainda que por meio dos próprios sistemas automatizados (artigo 19 do Decreto nº 9.936/2019) – nas e das tomadas de decisão algorítmicas.

Outras relevantes formas de *accountability* estão no treinamento, desenvolvimento e implementação dos algoritmos com dados representativos e participação plural de pessoas, visando à equidade algorítmica (Ferrari, 2023)<sup>11</sup> (*algorithmic fairness*). A garantia da pluralidade, na engenharia de dados, é crucial à redução dos potenciais efeitos discriminatórios dos algoritmos.

A implementação de *sandboxes* regulatórias é, também, uma alternativa promissora: trata-se “da estruturação de um ambiente regulatório experimental sob supervisão de uma autoridade pública” (Instituto de Tecnologia & Sociedade do Rio, 2024, p. 15). As *sandboxes* permitem a testagem de novas abordagens regulatórias, com um grupo pré-selecionado de especialistas (Instituto de Tecnologia & Sociedade do Rio, 2024, p. 15).

Outra alternativa interessante de *accountability* é o *design* centrado no usuário: alguns especialistas sugerem que deve ser assegurado efetivo grau de participação do indivíduo na configuração da decisão automatizada, “por exemplo, permitindo que ele afaste ou calibre a relevância de determinados resultados ou critérios de decisão” (Wimmer; Doneda, 2021, p. 398).

Especificamente no que diz respeito à responsabilização, é possível associá-la à redução de riscos, “em caso de negligência seguida da concretização dos riscos previamente identificados” (Instituto de Tecnologia & Sociedade do Rio, 2024, p. 15). Todavia, “o risco identificado pode ser alto demais, a ponto de se justificar a proibição de um determinado sistema da IA” (Instituto de Tecnologia & Sociedade do Rio, 2024, p. 15).

A *accountability* algorítmica, contudo, não vem livre de imensos desafios, pois cada qual das ferramentas citadas possui suas próprias limitações e desvantagens. A exemplo, mencionam-se as críticas no sentido de que: uma revisão humana ou explicação a *posteriori* da decisão algorítmica pode demorar a vir (Wimmer; Doneda, 2021); os relatórios de transparência não garantem cognoscibilidade; algumas interferências são prejudiciais à inovação tecnológica; não é possível prever todos os riscos das decisões automatizadas; e, ainda, tais mecanismos podem dar a impressão de que não há aplicações positivas de modelos matemáticos e de IA (Instituto de Tecnologia & Sociedade do Rio, 2024).

Não se desconhecem, portanto, as inúmeras críticas à *accountability* algorítmica, em especial, o recorrente argumento de que pode “colocar em risco o segredo comercial que protege sistemas algorítmicos no mercado de novas tecnologias” (Instituto de Tecnologia & Sociedade do Rio, 2024, p. 14). No entanto, as inúmeras violações de direitos humanos, ocasionadas por decisões automatizadas, justificam a construção urgente de um dever de *accountability* na matéria. O debate, aliado à educação emancipatória sobre a temática, deve incluir toda a sociedade (Ferrari, 2023).

Portanto, longe de ser uma questão meramente tecnocêntrica ou matemática, a discussão sobre os algoritmos envolve a todas e a todos, demandando “todo um aparato de governança que

<sup>11</sup> Segundo Ferrari (2023, p. 129), a preocupação é endereçada pela Resolução CNJ 332/2020, nos artigos 4º, 5º, e 7º.

permita identificar e combater os vieses”, “além de parâmetros para definir a supervisão necessária e adequada para cada caso específico” (Ferrari, 2023, p. 159).

A construção de um dever de *accountability*, na realidade latino-americana, deve resistir à tentação de assumir modelos regulatórios como paradigmáticos e essencialmente superiores, sobretudo advindos do norte-global. É igualmente importante compreender que os mecanismos de supervisão e controle não são auto-excludentes e, portanto, podem ser adotados de forma complementar, para criar um panorama regulatório individualizado e mais adequado a uma realidade específica. Por isso, propostas pretensamente universalistas de IA e de modelos algorítmicos devem ser rechaçadas, e a pluralidade de cosmovisões, considerada (Instituto de Tecnologia & Sociedade do Rio, 2024). Nesse aspecto, é importante mencionar que ganham força as abordagens decoloniais da IA – com incentivos a sistemas descentralizados (Soares, 2024), à interdisciplinaridade e à reestruturação da relação com a tecnologia digital (Hora, 2024).

Deve-se perguntar, ainda, em quais setores é possível a aplicação de cada uma das espécies algorítmicas (programadas, não programadas, supervisionadas, não supervisionadas ou de reforço, entre outras), sobretudo em atenção à realidade brasileira.

Diante da intensa concentração de poder nas mãos das *Big Techs*, quer-se pensar de que maneira os diálogos entre os diferentes planos protetivos dos direitos humanos (Fachin, 2021), especialmente no que diz respeito ao Sistema Interamericano de Direitos Humanos (SIDH), podem fortalecer os Estados no sentido de prevenir e combater violações de direitos humanos ocasionadas por tomadas de decisões algorítmicas.

Por isso, indaga-se: é possível, a partir da Convenção Americana sobre Direitos Humanos (CADH), extrair um dever de regular e supervisionar a *accountability* algorítmica?

## **O dever de regular *accountability* algorítmica, respeitá-la e supervisioná-la, sob a Convenção Americana sobre Direitos Humanos**

A Convenção Americana sobre Direitos Humanos (CADH), ratificada pelo Brasil em 1992, é o principal tratado internacional de direitos humanos do Sistema Interamericano de Direitos Humanos (SIDH) (Organização dos Estados Americanos, 1969). Composto por dois principais órgãos, a Comissão Interamericana de Direitos Humanos (Comissão IDH) e a Corte Interamericana de Direitos Humanos (Corte IDH), o SIDH é o mecanismo regional de proteção e promoção dos direitos humanos nas Américas.

A CADH estabelece, em seu artigo 2º, que os Estados Partes da Convenção devem adotar disposições de direito interno para efetivar os direitos e as liberdades nela previstos, bem como verificar se tais disposições, caso existentes, estão de acordo com o *corpus iuris* interamericano (dever de caráter positivo). Caso averiguado que as normas estão em desacordo com os parâmetros do SIDH, cabe ao Estado o dever de derogá-las ou revogá-las (dever de caráter negativo). Busca-se, a partir do mencionado dispositivo, garantir efetividade material à CADH (Garbin, 2024).

A identificada ausência de regulação – de suficiente densidade – acerca dos sistemas algorítmicos, pormenorizada nos pontos anteriores, viola, pois, o artigo 2º da CADH. Os exemplos trazidos ao longo deste trabalho evidenciam que as lacunas normativas, quanto à *accountability* algorítmica, levam a violações de outros direitos humanos, também previstos na Convenção, como, em conjunto com o artigo 1º (obrigação de respeito), aos artigos 5º (direito à integridade pessoal), 11 (proteção da honra e da dignidade), 13 (liberdade de pensamento e de expressão), 26 (desenvolvimento progressivo), entre outros.

Exemplo nítido de como a violação ao artigo 2º, conjuntamente com o artigo 1º, pode implicar afronta ao artigo 26 da CADH (desenvolvimento progressivo) foi a tentativa de implementação de robôs como estratégia de automatização na análise inicial de pedidos de benefícios ao Instituto Nacional do Seguro Social (INSS).

O Tribunal de Contas da União (TCU), em recente auditoria, identificou que a resposta padrão dos algoritmos ampliou significativamente o percentual de negativas, sem as devidas explicações dos motivos. O TCU também constatou a ausência de padrões mínimos de transparência e indicou que a referida tecnologia, sem as adaptações necessárias, levou à excessiva demora na finalização dos processos da seguridade social (Nem [...], 2023).

No caso em menção, fica patente a violação ao direito humano à seguridade social, direito autônomo protegido pelo artigo 26 da Convenção (Corte Interamericana de Direitos Humanos, 2021), decorrente da ausência de disposições internas – suficientes – sobre *accountability* algorítmica.

Além disso, a Corte IDH já estabeleceu que os Estados também possuem a obrigação de supervisionar, tanto no ponto de vista da prevenção e fiscalização quanto no da sanção, atividades empresariais (Volz, 2024). Por isso, em conjunto com a necessária observância ao artigo 2º da CADH, os Estados devem estabelecer obrigações vinculantes às empresas e fiscalizar, adequadamente, as atividades desenvolvidas, sob pena de responsabilização estatal por omissão (Corte Interamericana de Direitos Humanos, 2006; Volz, 2024).

A Comissão IDH já tem expressado, há alguns anos, sérias preocupações com os efeitos das decisões algorítmicas. No Relatório Empresas e Direitos Humanos (Organização dos Estados Americanos, 2019), destacou a necessidade de se adotar mecanismos de transparência, controle e equilíbrios adequados – em suma, de *accountability* algorítmica –, para evitar discriminações, manipulação informativa, violações à privacidade, aos dados pessoais, ao anonimato, entre outras afrontas aos direitos humanos.

Ainda, recentemente, a Corte IDH reconheceu a *autodeterminação informativa* – ou seja, o direito que cada pessoa tem de controlar os próprios dados pessoais, “decidindo como e por quem esses dados podem ser coletados, usados e divulgados” (Fachin; Ramos, 2024, p. 21) – como um direito autônomo, no Caso Membros da Corporação Coletivo de Advogados “José Alvear Restrepo” Vs. Colômbia.

Por fim, cabe aos próprios Estados, igualmente, a obrigação de respeitar as regulações sobre *accountability* algorítmica: nos sistemas de justiça – que, cada vez mais, têm utilizado IA –, nas entidades da Administração Pública direta e indireta, nos setores da saúde, transporte, segurança, gestão de serviços públicos, entre outros.

Portanto, evidencia-se que o dever estatal de regular e supervisionar a *accountability* algorítmica pode ser extraído da CADH, impondo-se aos Estados Partes, incluindo-se o Brasil, o seu urgente cumprimento, sob pena de responsabilização no plano internacional e, mais que isso, de se permitir o acentuamento das discriminações e ainda maiores violações aos direitos humanos, especialmente de populações em maior vulnerabilidade social.

## Conclusão

O presente artigo, a partir da metodologia hipotético-dedutiva, explicou os aspectos básicos dos algoritmos, escancarou as potenciais ameaças dos sistemas de decisões automatizadas às democracias e aos direitos humanos e ressaltou a urgência de se estabelecer, com suficiente densidade, a *accountability* algorítmica no âmbito do ordenamento jurídico interno.

Evidenciou-se, neste trabalho, como a crescente implementação de decisões automatizadas impõe desafios éticos, jurídicos e políticos sem precedentes, uma vez que a opacidade desses sistemas, aliada à arquitetura virtual assimétrica, perpetua desigualdades, reforça o racismo, a violência de gênero e outras discriminações, influencia em decisões políticas, viola direitos humanos e ameaça a democracia constitucional. Criticou-se, também, a noção de neutralidade algorítmica, revelando que as decisões automatizadas são moldadas por dados enviesados, que refletem e amplificam estruturas de poder.

Nesse contexto, a *accountability* algorítmica emerge como um imperativo regulatório, fundamental à garantia de transparência e supervisão das decisões automatizadas, que se desdobra em distintas dimensões, abrangendo a necessidade de regulamentação estatal, mecanismos de auditoria, avaliações de impacto algorítmico, divulgação periódica de relatórios de transparência, entre outros.

Ao final, a partir da perspectiva multinível dos direitos humanos, confirmou-se a hipótese da pesquisa: é possível extrair, da Convenção Americana sobre Direitos Humanos (CADH), o dever do Estado brasileiro de adotar disposições internas sobre *accountability* algorítmica (em observância ao artigo 2º da CADH), bem como de respeitá-la (artigo 1º da CADH) e supervisioná-la. A inobservância desses deveres, inclusive a omissão em cumpri-los, impõe a responsabilização do Brasil no âmbito internacional, viola direitos humanos e a democracia constitucional multinível. O direito não deve se esquivar desse debate, fundamentalmente político.

Diante da crescente automação dos processos decisórios em todos os âmbitos da vida, a *accountability* deve ser tratada como uma prioridade na agenda política e jurídica. Estados, organismos internacionais, empresas, comunidade científica e cidadãos precisam estar comprometidos com uma abordagem regulatória eficaz e interdisciplinar, com respeito às singularidades da *internet*. No contexto brasileiro, há de se ficar vigilante quanto ao PL nº 2.338 de 2023, que promete regular a inteligência artificial no país. Ao fazê-lo, o Brasil acena, ainda que lentamente, ao cumprimento de suas obrigações internacionais de direitos humanos em matéria algorítmica.

## Referências

- Algorithmic Transparency Recording Standard Hub. GOV.UK, 5 jan. 2023. Available from: <https://www.gov.uk/government/collections/algorithmic-transparency-recording-standard-hub>. Cited: 2025 May 23.
- Bezerra, A. C. B.; Costa, C. M. Pele negra, algoritmos brancos: informação e racismo nas redes sociotécnicas. *Liinc em Revista*, v. 18, n. 2, e6043, 2022. Doi: <https://doi.org/10.18617/liinc.v18i2.6043>.
- Branco, S.; Teffé, C. S. (coord.). *Inteligência Artificial e Big Data: diálogos da pós-graduação em Direito Digital*. Rio de Janeiro: Instituto de Tecnologia e Sociedade do Rio de Janeiro, 2022. 327 p. Disponível em: [https://itsrio.org/wp-content/uploads/2022/10/Livro\\_P%C3%B3s-ITS-UERJ\\_COMPLETO\\_v3.pdf](https://itsrio.org/wp-content/uploads/2022/10/Livro_P%C3%B3s-ITS-UERJ_COMPLETO_v3.pdf). Acesso em: 17 abr. 2024.
- Burrell, B. How the machine ‘thinks’: understanding opacity in machine learning algorithms. *Big Data & Society*, v. 3, n. 1, p. 1-12, 2016.
- Carrera, F.; Carvalho, D. Algoritmos racistas: a hiper-ritualização da solidão da mulher negra em bancos de imagens digitais. *Galáxia*, n. 43, p. 99-114, 2020. Doi: <http://dx.doi.org/10.1590/1982-25532020141614>.
- Coalizão Direitos na Rede. Projeto de Lei n. 2338/2023. *Direitos na Rede*, 18 jun. 2024. Disponível em: <https://direitosnarede.org.br/2024/06/18/projeto-de-lei-n-2338-2023/>. Acesso em: 16 maio 2025.
- Conselho Nacional de justiça (Brasil). *Protocolo para Julgamento com Perspectiva de Gênero*. Brasília: CNJ, 2021. Disponível em: <https://www.cnj.jus.br/wp-content/uploads/2021/10/protocolo-para-julgamento-com-perspectiva-de-genero-cnj-24-03-2022.pdf>. Acesso em: 16 maio 2025.

- Corte Interamericana de Derechos Humanos. *Caso Vera Rojas e Outros Vs. Chile*. Exceções preliminares, Mérito, Reparações e Custas. Sentença de 1 de outubro de 2021. San José: CORTE IDH, 2021. Disponível em: [https://www.corteidh.or.cr/docs/casos/articulos/seriec\\_439\\_esp.pdf](https://www.corteidh.or.cr/docs/casos/articulos/seriec_439_esp.pdf). Acesso em: 2 jun. 2025.
- Corte Interamericana de Derechos Humanos. *Caso Ximenes Lopes Vs. Brasil*. Mérito, Reparações e Custas. Sentença de 4 de julho de 2006. Disponível em: [https://www.corteidh.or.cr/docs/casos/articulos/seriec\\_149\\_por.pdf](https://www.corteidh.or.cr/docs/casos/articulos/seriec_149_por.pdf). Acesso: 10 dez. 2025.
- Dastin, J. Insight – Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters*, 11 out. 2018. Available from: <https://www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG/>. Cited: 2025 Dec 11.
- European Approach to Artificial Intelligence. *European Commission*, [202-?]. Disponível em: <https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence>. Acesso em: 23 maio 2025.
- Fachin, M. Constitucionalismo multinível: diálogos e(m) direitos humanos. *Revista Ibérica do Direito*, v. 1, n. 1, p. 53-68, 2021. Disponível em: <https://www.revistaibericadodireito.pt/index.php/capa/article/view/26>. Acesso em: 17 abr. 2024.
- Fachin, M.; Ramos, C. M. V. Direito à autodeterminação informativa. In: Fachin, M. G.; Câmara, F. H. (coord.). *Dicionário Interamericano de Direitos Humanos*. Volume 2: direitos e deveres. Rio de Janeiro: Lumen Juris, 2024. p. 21-23.
- Ferrari, I. *Accountability* de algoritmos: a falácia do acesso ao código e caminhos para uma explicabilidade efetiva. *Inteligência Artificial: 3º Grupo de Pesquisa do ITS*. Rio de Janeiro: ITS - Instituto de Tecnologia e Sociedade do Rio, 2018. f. 1-18. Disponível em: <https://itsrio.org/wp-content/uploads/2019/03/Isabela-Ferrari.pdf>. Acesso em: 23 maio 2025.
- Ferrari, I. *Discriminação algorítmica e poder judiciário*: limites à adoção de sistemas de decisões algorítmicas no judiciário brasileiro. Florianópolis: Emais Editora, 2023.
- Garbin, G. A. L. Dever de adotar disposições de direito interno. In: Fachin, M. G.; Câmara, F. H. (coord.). *Dicionário Interamericano de Direitos Humanos*. Volume 2: direitos e deveres. Rio de Janeiro: Lumen Juris, 2024. p. 341-342.
- Hora, N. Da neutralidade à IA decolonial. *ComCiência*, 15 jun. 2024. Disponível em: <https://www.comciencia.br/da-neutralidade-a-ia-decolonial/>. Acesso em: 23 maio 2025.
- Instituto de Tecnologia & Sociedade do Rio. *Como regular a inteligência artificial?* Expandindo horizontes de análise para além da União Europeia. Rio de Janeiro, 2024. 26 p. Disponível em: [https://itsrio.org/wp-content/uploads/2016/12/20240404\\_Relatorio\\_Microsoft\\_Como-Regular-IA\\_3.pdf](https://itsrio.org/wp-content/uploads/2016/12/20240404_Relatorio_Microsoft_Como-Regular-IA_3.pdf). Acesso em: 24 abr. 2024.
- Köchling, A.; Wehner, M. C. Discriminated by an algorithm: a systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HR development. *Business Research, Düsseldorf*, v. 13, p. 795-848, 2020. Doi: <https://doi.org/10.1007/s40685-020-00134-w>. Acesso em: 24 abr. 2024.
- Kroll, J. A. et al. Accountable Algorithms. *University of Pennsylvania Law Review*, v. 165, n. 2765268, p. 1-66, 2016. Disponível em: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2765268](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2765268). Acesso em: 24 abr. 2025.
- Lemos, R. China cria a primeira regulação para algoritmos. *Folha de São Paulo*, 8 set. 2021. Artigos. Disponível em: <https://itsrio.org/pt/artigos/china-cria-a-primeira-regulacao-para-algoritmos/>. Acesso em: 23 maio 2025.
- Lopes, A. J.; Paola, E. Precisamos conter os riscos da inteligência artificial, diz Lula. In: Poder360. *Governo*, 23 maio. 2024. Disponível em: <https://www.poder360.com.br/governo/precisamos-conter-os-riscos-da-inteligencia-artificial-diz-lula/>. Acesso em: 16 maio 2025.
- Mattos Filho. Inteligência artificial: avaliação de impacto algorítmico. *Notícias e Negócios*, 26 set. 2023. Disponível em: <https://www.mattosfilho.com.br/unico/inteligencia-artificial-avaliacao-impacto-algoritmico/>. Acesso em: 16 maio 2025.
- Mota, R. K. B. et al. Concepções teóricas sobre *accountability*. In: Congresso UFU de Contabilidade, 4., 2021, Uberlândia. Uberlândia: UFU, 2021. f. 1-19. Disponível em: <https://eventos.ufu.br/sites/eventos.ufu.br/files/documentos/contufu2021.completo0078.pdf>. Acesso em: 23 maio 2025.
- Nem tudo são flores no mundo da Inteligência Artificial. *Estadão*, 7 ago. 2023. Disponível em: <https://www.estadao.com.br/politica/gestao-politica-e-sociedade/nem-tudo-sao-flores-no-mundo-da-inteligencia-artificial/>. Acesso em: 23 maio 2025.

Noble, S. U. *Algorithms of Oppression: how search engines reinforce racism*. [S. l.]: NYU Press, 2018.

O'Neil, C. *Algoritmos de destruição em massa: como o big data aumenta a desigualdade e ameaça a democracia*. Santo André: Rua do Sabão, 2020.

Ordóñez, V.; Dunn, T.; Noll, E. OpenAI CEO Sam Altman says AI will reshape society, acknowledges risks: 'A little bit scared of this'. *ABC News Technology*, 16 mar. 2023. Disponível em: <https://abcnews.go.com/Technology/openai-ceo-sam-altman-ai-reshape-society-acknowledges/story?id=97897122>. Acesso em: 23 maio 2025.

Organização dos Estados Americanos. *Comissão Interamericana de Direitos Humanos. Empresas e Direitos Humanos: padrões interamericanos*. Washington, D.C.: OEA, 2019. Disponível em: <https://www.oas.org/pt/cidh/relatorios/pdfs/empresas%20e%20direitos.pdf>. Acesso em: 2 jun. 2025.

Organização dos Estados Americanos. *Convenção Americana sobre Direitos Humanos (Pacto de San José da Costa Rica)*. San José: Comissão Interamericana de Direitos Humanos, 1969. Disponível em: [https://www.cidh.org/basicos/portugues/c.convencao\\_americana.htm](https://www.cidh.org/basicos/portugues/c.convencao_americana.htm). Acesso em: 15 dez. 2025.

Organization for Economic Co-operation and Development. *Global trends in government innovation 2023*. OECD Public Governance Reviews. Paris: OECD Publishing, 2023. Doi: <https://doi.org/10.1787/0655b570-en>. Acesso em: 16 maio 2025.

Parra, H. Z. M.; Abdo, A. H. Tendências democráticas e autoritárias, arquiteturas distribuídas e centralizadas. In: Bragatto, R. C.; Sampaio, R. C.; Silva, S. P. (org.). *Democracia digital, comunicação política e redes: teoria e prática*. Rio de Janeiro: Folio Digital: Letra e Imagem, 2016. cap. 5, p. 137-159.

Rached, D. H. *Accountability* e desenho institucional: um ponto cego no direito público brasileiro. *Revista Brasileira de Políticas Públicas*, v. 11, n. 1, p. 190-209, 2021. Doi: <https://doi.org/10.5102/rbpp.v11i1.6745>.

Regulação Algorítmica. Entrevistado: Karen Young. *Digilabour*, 14 mar. 2020. Interviews. Disponível em: <https://digilabour.com.br/pt/regulacao-algoritmica-entrevista-com-karen-young/>. Acesso em: 23 maio 2025.

Ribeiro, E. Decifrando a tecnologia do reconhecimento facial. *Medium: Human Stories & Ideas*, 17 jan. 2023. Disponível em: [https://medium.com/@elieser\\_ribeiro/decifrando-a-tecnologia-do-reconhecimento-facial-27ee45eb6ada](https://medium.com/@elieser_ribeiro/decifrando-a-tecnologia-do-reconhecimento-facial-27ee45eb6ada). Acesso em: 17 abr. 2024.

Saldanha, P. M. Governança, *accountability* e avaliação de impacto algorítmico. *Conjur – Consultor Jurídico*, 4 jun. 2024. Disponível em: <https://www.conjur.com.br/2024-jun-04/governanca-accountability-e-avaliacao-de-impacto-algoritmico/>. Acesso em: 23 maio 2025.

Sfredo, M.; Lopes, R. Um “novo Estado”? Mais ricas do que muitos países, empresas gigantes da tecnologia são cobradas por suas práticas: Google, Facebook, Apple, Amazon e Microsoft têm sido chamadas a explicar suas atuações nem sempre saudáveis no ambiente democrático. *GZH*, 18 dez. 2020. Disponível em: <https://gauchazh.clicrbs.com.br/tecnologia/noticia/2020/12/um-novo-estado-mais-ricas-do-que-muitos-paises-empresas-gigantes-da-tecnologia-sao-cobradas-por-suas-praticas-ckiduceo70001019woewvaey2.html>. Acesso em: 16 maio 2025.

Silva, T. Visão computacional e racismo algorítmico: branquitude e opacidade no aprendizado de máquina. *Revista da Associação Brasileira de Pesquisadores/as Negros/as*, v. 12, n. 31, p. 428-448, 2020. Doi: <https://doi.org/10.31418/2177-2770.2020>.

Soares, M. Uma IA decolonial é possível? *Desinformante: informação confiável sobre desinformação*, 26 jun. 2024. Disponível em: <https://desinformante.com.br/ia-decolonial/>. Acesso em: 23 maio 2025.

Volz, M. B. Dever de supervisionar atividades empresariais. In: Fachin, M. G.; Câmara, F. H. (coord.). *Dicionário Interamericano de Direitos Humanos*. Volume 2: direitos e deveres. Rio de Janeiro: Lumen Juris, 2024. p. 425-427.

Wimmer, M.; Doneda, D. “Falhas de IA” e a intervenção humana em decisões automatizadas: parâmetros para a legitimação pela humanização. *Revista Direito Público*, v. 18, n. 100, p. 374-404, 2021. Doi: <https://doi.org/10.11117/rdp.v18i100.6119>. Acesso em: 25 maio 2025.