Journal of Legal Sciences
Pensar
e-ISSN: 2317-2150
Universidade de Fortaleza

https://doi.org/10.5020/2317-2150.2026.17037

Decisões automatizadas e riscos algorítmicos: por uma interpretação da discriminação algorítmica à luz dos direitos fundamentais

Automated decisions and algorithmic risks: towards an interpretation of algorithmic discrimination in light of fundamental rights

Decisiones automatizadas y riesgos algorítmicos: por una interpretación de la discriminación algorítmica a la luz de los derechos fundamentales

Lucas Moreschi Paulo* https://orcid.org/0000-0003-4583-4853, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brasil.

Editorial

Histórico do Artigo
Recebido: 22/06/2026
Aceito: 21/08/2026

Eixo Temático 3: Direito, Tecnologia e Sociedade em Transformação

Editores-chefes
Katherinne de Macêdo Maciel Mihaliuc
Universidade de Fortaleza, Fortaleza, Ceará, Brasil
katherinne@unifor.br

Sidney Soares Filho
Universidade de Fortaleza, Fortaleza, Ceará, Brasil
sidney@unifor.br

Editor Responsável
Sidney Soares Filho
Universidade de Fortaleza, Fortaleza, Ceará, Brasil
sidney@unifor.br

Autor
Lucas Moreschi Paulo
lucasmoreschipaulo@gmail.com
Contribuição: conceituação; metodologia; investigação; análise formal; curadoria bibliográfica; redação do rascunho original; redação, revisão e edição; validação final do manuscrito.

Como citar:
PAULO, Lucas Moreschi. Decisões automatizadas e riscos algorítmicos: por uma interpretação da discriminação algorítmica à luz dos direitos fundamentais. **Pensar – Revista de Ciências Jurídicas**, Fortaleza, v. 31, e17037, 2026. DOI: https://doi.org/10.5020/2317-2150.2026.17037

Declaração de disponibilidade de dados

A Pensar – Revista de Ciências Jurídicas adota práticas de Ciência Aberta e disponibiliza, junto à presente publicação, a Declaração de Disponibilidade de Dados (Formulário Pensar Data) preenchida e assinada pelos autores, a qual contém informações sobre a natureza do artigo e a eventual existência de dados complementares. O documento pode ser consultado como arquivo suplementar neste site.

Resumo

O artigo examina a discriminação algorítmica como problema jurídico-constitucional produzido pela convergência entre decisões automatizadas, opacidade técnica e desigualdades estruturais. Parte-se da hipótese de que a discriminação algorítmica não deve ser interpretada apenas como defeito técnico do modelo, erro estatístico ou uso indevido de dados pessoais, mas também como forma contemporânea de afetação de direitos fundamentais, quando sistemas automatizados classificam, ranqueiam, recomendam ou excluem pessoas em domínios sensíveis da vida social. A pesquisa adota método lógico-dedutivo, com revisão bibliográfica e análise normativa, articulando direito antidiscriminatório, proteção de dados, proporcionalidade e teoria dos direitos fundamentais. O argumento incorpora o modelo decisório epistêmico-ponderativo desenvolvido em pesquisa doutoral, no qual a intensidade da intervenção, o peso abstrato dos direitos em colisão e a certeza das premissas técnicas orientam a gradação A1-A3 de risco algorítmico. Sustenta-se que a resposta constitucional adequada deve deslocar o debate da simples transparência formal para um modelo de justificabilidade algorítmica, no qual abertura informacional, explicabilidade, auditabilidade, supervisão humana significativa e distribuição dinâmica dos ônus probatórios sejam calibradas conforme o risco e o impacto sobre a pessoa afetada ou sobre grupos em vulnerabilidade. Conclui-se que, quanto maior a intensidade da interferência e menor a certeza das premissas técnicas, maior deve ser o dever de demonstração, de correção e de prevenção imposto ao controlador, ao desenvolvedor, ao utilizador institucional e ao Estado.

Palavras-chave: decisões automatizadas; discriminação algorítmica; direitos fundamentais; explicabilidade; proporcionalidade.

Abstract

This article examines algorithmic discrimination as a constitutional-legal problem produced by the convergence of automated decisions, technical opacity and structural inequalities. It starts from the hypothesis that algorithmic discrimination should not be interpreted merely as a technical defect, statistical error or improper use of personal data, but as a contemporary form of fundamental-rights interference whenever automated systems classify, rank, recommend or exclude persons in sensitive domains of social life. The research adopts a logical-deductive method, with bibliographical review and normative analysis, articulating anti-discrimination law, data protection, proportionality and fundamental-rights theory. The argument incorporates an epistemic-proportional decision-making model developed in doctoral research, in which the intensity of interference, the abstract weight of colliding rights and the certainty of technical premises guide the A1-A3 scale of algorithmic risk. It argues that the adequate constitutional response must move beyond mere formal transparency towards a model of algorithmic justifiability, in which informational openness, explainability, auditability, meaningful human oversight and dynamic distribution of evidentiary burdens are calibrated according to risk and impact. It concludes that the greater the intensity of the interference and the lower the certainty of the technical premises, the stronger the duty of demonstration, correction and prevention imposed on controllers, developers, institutional users and the State.

Keywords: *automated decisions; algorithmic discrimination; fundamental rights; explainability; proportionality.*

Resumen

El artículo examina la discriminación algorítmica como problema jurídico-constitucional producido por la convergencia entre decisiones automatizadas, opacidad técnica y desigualdades estructurales. Se parte de la hipótesis de que la discriminación algorítmica no debe ser interpretada solo como un defecto técnico del modelo, un error estadístico o el uso indebido de datos personales, sino también como una forma contemporánea de afectación de derechos fundamentales, cuando los sistemas automatizados clasifican, jerarquizan, recomiendan o excluyen personas en dominios sensibles de la vida social. La investigación adopta un método lógico-deductivo, con revisión bibliográfica y análisis

* Doutor em Direito pela Universidade de Santa Cruz do Sul. Mestre em Direito pela Fundação Escola Superior do Ministério Público. Pós-doutorando em Direito pela Universidade Federal do Rio Grande do Sul. Advogado, professor e pesquisador em teoria do direito, jurisdição constitucional, direitos fundamentais e inteligência artificial.

normativo, articulando derecho antidiscriminatorio, protección de datos, proporcionalidad y teoría de los derechos fundamentales. El argumento incorpora el modelo decisorio epistémico-ponderativo desarrollado en la investigación doctoral, en el cual la intensidad de la intervención, el peso abstracto de los derechos en colisión y la certeza de las premisas técnicas orientan la gradación A1-A3 de riesgo algorítmico. Se sostiene que la respuesta constitucional adecuada debe desplazar el debate de la simple transparencia formal hacia un modelo de justificabilidad algorítmica, en el cual la apertura informacional, la explicabilidad, la auditabilidad, la supervisión humana significativa y la distribución dinámica de la carga probatoria sean calibradas conforme al riesgo y al impacto sobre la persona afectada o sobre grupos en vulnerabilidad. Se concluye que, a mayor intensidad de la interferencia y menor certeza de las premisas técnicas, mayor debe ser el deber de demostración, de corrección y de prevención impuesto al controlador, al desarrollador, al usuario institucional y al Estado.

Palabras clave: *decisiones automatizadas; discriminación algorítmica; derechos fundamentales; explicabilidad; proporcionalidad.*

1 Introdução

A decisão automatizada deixou de ser uma hipótese periférica da vida digital para converter-se em uma das formas ordinárias de distribuição de oportunidades, de riscos, de suspeitas e de expectativas sociais. Os sistemas algorítmicos classificam consumidores, ordenam candidatos a emprego, calculam riscos de crédito, modulam preços, direcionam publicidade, recomendam conteúdos, priorizam fiscalizações, organizam filas administrativas, apoiam diagnósticos, indicam padrões de fraude e, progressivamente, influenciam também a gestão da atividade jurisdicional. O problema jurídico não está, por si só, na utilização de técnicas computacionais para lidar com grandes volumes de dados, pois seria ingênuo negar a relevância pública da eficiência, da escala e da capacidade de identificação de padrões em sociedades marcadas por complexidade informacional. O problema surge quando essa eficiência se apresenta como substituto da razão pública e quando a regularidade estatística passa a funcionar, na prática, como critério silencioso de inclusão ou exclusão de pessoas.

A discriminação algorítmica deve ser compreendida nesse ponto de contato entre técnica e Constituição. Ela não se reduz ao preconceito subjetivo do programador, à intenção discriminatória do agente econômico ou ao erro isolado de uma linha de código. Ainda que todas essas hipóteses possam existir, o núcleo do fenômeno é mais profundo: sistemas automatizados podem reproduzir, amplificar ou ocultar padrões desiguais já inscritos nos dados, nas categorias de classificação, nas métricas de desempenho, nos objetivos de otimização e nas formas institucionais de uso. A tecnologia, nesse sentido, não cria a desigualdade a partir do nada; em vez disso, reconfigura esse fenômeno, tornando-o mais veloz, menos visível e, muitas vezes, menos contestável.

O tema ganha especial densidade no constitucionalismo brasileiro porque a Constituição Federal de 1988 (Brasil, [2026a]) não trata a igualdade como mera proibição de diferenciações arbitrárias, mas como compromisso material de construção de uma sociedade livre, justa e solidária. A dignidade da pessoa humana, a igualdade, a liberdade, a privacidade, a proteção de dados pessoais, o devido processo legal, o contraditório e a vedação de discriminações injustificadas compõem uma gramática normativa que impede a submissão do cidadão a classificações opacas, sobretudo quando, delas, decorrem efeitos relevantes sobre crédito, trabalho, educação, saúde, segurança pública, assistência social ou acesso à justiça. A Lei Geral de Proteção de Dados Pessoais (LGPD) reforça esse desenho ao reconhecer, entre outros pontos, os princípios de finalidade, adequação, necessidade, transparência, segurança, prevenção, não discriminação e responsabilização, além de prever o direito de solicitar revisão de decisões tomadas unicamente com base em tratamento automatizado de dados pessoais que afetem interesses do titular (Brasil, [2026b]).

A experiência europeia, embora não possa ser transplantada sem mediações, confirma a centralidade do problema. O Regulamento Geral de Proteção de Dados estabelece proteção específica contra decisões baseadas exclusivamente em tratamento automatizado, inclusive perfilização, quando produzirem efeitos jurídicos ou afetarem, significativamente, a pessoa (União Europeia, 2016). O Regulamento Europeu de Inteligência Artificial, por sua vez, consolida um modelo regulatório orientado por risco, com deveres mais intensos para sistemas de alto risco, especialmente em áreas como educação, emprego, acesso a serviços essenciais, migração, justiça e aplicação da lei (União Europeia, 2024). A direção é relevante: quanto maior o risco sobre direitos, maior deve ser o ônus da governança, da documentação, da supervisão, da transparência e do controle.

A pergunta que orienta este artigo pode ser formulada nos seguintes termos: de que modo a discriminação algorítmica deve ser interpretada à luz dos direitos fundamentais, quando decisões automatizadas produzem efeitos relevantes sobre pessoas e grupos socialmente vulnerabilizados? A hipótese sustentada é a de que a resposta adequada não se encontra nem na rejeição abstrata da tecnologia nem na confiança acrítica em sua suposta neutralidade. A interpretação constitucional deve construir um modelo de justificabilidade algorítmica, no qual a

validade do uso de sistemas automatizados dependa da possibilidade de reconstrução pública das razões técnicas e normativas que sustentaram a decisão, da demonstração de que os riscos foram previamente avaliados, da existência de supervisão humana significativa e da abertura proporcional das informações necessárias à contestação.

A formulação deste estudo incorpora, de modo central, a matriz desenvolvida por Paulo (2026), em pesquisa doutoral sobre discriminação algorítmica, proteção de grupos em vulnerabilidade e dever estatal de proteção à luz da lei material da ponderação. O ganho desse deslocamento está em evitar que a análise permaneça apenas no plano da transparência ou da proteção de dados: a decisão automatizada passa a ser examinada como prática de afetação jusfundamental, cuja legitimidade depende do grau de risco, da intensidade da interferência, da qualidade epistêmica das premissas técnicas e da possibilidade real de contestação pelos afetados.

A pesquisa adota o método lógico-dedutivo, com revisão bibliográfica e análise normativa. O percurso desenvolve-se em quatro movimentos. Primeiro, examinam-se as decisões automatizadas e os riscos algorítmicos, mostrando como a automação leva a decisão para uma arquitetura sociotécnica, composta por dados, modelos, objetivos de otimização e usos institucionais. Segundo, reconstrói-se a passagem da discriminação estrutural à discriminação algorítmica, para demonstrar que a neutralidade estatística pode funcionar como linguagem contemporânea de desigualdades materiais. Terceiro, propõe-se uma interpretação jusfundamental da discriminação algorítmica, a partir da dignidade, da igualdade, da autodeterminação informativa e do devido processo. Quarto, apresentam-se critérios de controle: explicabilidade suficiente, auditabilidade, abertura informacional proporcional ao impacto, supervisão humana significativa e inversão dinâmica do ônus informacional.

O autor declara que utilizou a ferramenta *ChatGPT* para apoio à organização textual, à estruturação preliminar, à revisão linguística e à padronização formal do manuscrito. A responsabilidade pelo conteúdo final, pela precisão das informações apresentadas, pela seleção das fontes, pela originalidade do texto e pelas conclusões defendidas é integralmente humana.

2 Decisões automatizadas e risco algorítmico: da eficiência operacional à afetação de direitos

A expressão “decisão automatizada” costuma sugerir uma imagem simples: uma máquina recebe dados, processa informações e entrega um resultado. Essa descrição, embora funcionalmente correta em nível superficial, é insuficiente para o Direito. Uma decisão automatizada não é apenas o resultado final apresentado ao usuário, ao consumidor, ao jurisdicionado ou ao administrado, mas é também o produto de uma cadeia de escolhas humanas, técnicas e institucionais, que começa na definição do problema, passa pela seleção dos dados, pela construção das variáveis, pela escolha do modelo, pela fixação da métrica de sucesso, pela parametrização do sistema, pela definição de limiares decisórios e termina no modo como a organização utiliza o resultado. A decisão, portanto, é automatizada no momento de sua execução, mas não é puramente maquínica em sua origem.

Essa constatação tem relevância jurídica imediata. Se a decisão automatizada é uma arquitetura sociotécnica, a responsabilidade por seus efeitos não pode ser deslocada para a opacidade do sistema, como se a técnica fosse um sujeito autônomo. A automação não extingue a imputação; ao contrário, estabelece que se reconstrua a cadeia de produção do resultado para identificar quem definiu o objetivo, quem selecionou os dados, quem treinou ou contratou o modelo, quem validou o sistema, quem o colocou em funcionamento, quem monitorou os resultados e quem se beneficiou da decisão. A opacidade algorítmica, nesse ponto, não pode funcionar como zona de irresponsabilidade.

Nesse cenário, o risco algorítmico nasce da distância entre a aparência de neutralidade do cálculo e a densidade normativa dos efeitos que dele decorrem. Assim, os sistemas automatizados podem parecer objetivos, porque expressam resultados em números, *scores*, *rankings* ou classificações. Essa forma de apresentação, contudo, não elimina a presença de pressupostos valorativos. As decisões sobre quais dados importam, quais variáveis são legítimas, qual erro é tolerável, qual grupo será mais exposto a falsos positivos, qual custo será priorizado e qual finalidade institucional deve ser maximizada são escolhas carregadas de normatividade. A técnica não está fora do Direito; apenas conduz parte da normatividade para camadas menos acessíveis à linguagem jurídica ordinária.

O Direito, por isso, não deve se contentar com a pergunta sobre se o sistema funciona em termos médios. A questão constitucional decisiva é outra: funciona para quem, contra quem, com quais custos, com que grau de contestabilidade e sob qual justificativa pública? Um modelo de crédito pode apresentar alta acurácia geral e, ainda assim, excluir desproporcionalmente moradores de determinadas regiões; um sistema de recrutamento pode

reduzir custos empresariais e, ao mesmo tempo, penalizar trajetórias profissionais marcadas por maternidade ou cuidado familiar; uma ferramenta de policiamento preditivo pode identificar áreas com maior incidência registrada de crimes e, simultaneamente, retroalimentar padrões históricos de seletividade policial. A eficiência média não responde ao problema da igualdade material.

O risco também se agrava pela assimetria informacional. A pessoa afetada por uma decisão automatizada, normalmente, não conhece os dados utilizados, as inferências produzidas, os critérios de classificação, as bases de comparação, as margens de erro, as variáveis de aproximação ou os mecanismos de contestação disponíveis. Muitas vezes, sequer sabe que foi objeto de tratamento automatizado. Essa assimetria transforma o contraditório em promessa vazia: não se contesta adequadamente aquilo que não se conhece; não se refuta uma inferência inacessível e não se demonstra discriminação quando a prova essencial está concentrada nas mãos de quem controla o sistema.

A literatura sobre opacidade algorítmica demonstra que a falta de compreensão dos modelos pode decorrer de diferentes fontes. Burrell (2016) identifica formas de opacidade ligadas ao sigilo corporativo, à distância técnica entre especialistas e leigos e à própria complexidade de modelos de aprendizagem de máquina, cuja lógica interna nem sempre é facilmente traduzível em linguagem natural. Pasquale (2015), por sua vez, descreve a sociedade da caixa-preta como um ambiente no qual decisões relevantes são tomadas por sistemas inacessíveis ao escrutínio dos afetados, do público e, por vezes, do próprio Estado. O ponto em comum dessas análises reside no fato de que a opacidade não é apenas obstáculo cognitivo, mas é também problema de poder.

A noção de risco algorítmico, portanto, deve ser compreendida em sentido ampliado. Há riscos técnicos, como baixa qualidade dos dados, sobreajuste, erro de classificação, instabilidade do modelo, ausência de validação externa e degradação de desempenho ao longo do tempo. Também há riscos institucionais, relacionados ao uso acrítico dos resultados, à falta de supervisão humana, à ausência de canais de revisão e à delegação informal de autoridade decisória ao sistema, bem como riscos jurídicos, ligados à violação de direitos fundamentais, à discriminação indireta, à falta de motivação e à impossibilidade de controle. Ademais, há riscos democráticos, quando sistemas automatizados passam a influenciar políticas públicas ou decisões estatais sem publicidade suficiente para permitir deliberação, fiscalização e responsabilização.

Nesse quadro, a técnica regulatória baseada em risco tem mérito por reconhecer que nem todo sistema automatizado requer o mesmo grau de controle. Um algoritmo que organiza documentos internos não produz a mesma intensidade de interferência de um sistema que decide acesso a benefício social, ao crédito habitacional, à vaga de emprego ou à liberdade provisória. Dessa forma, a proporcionalidade do controle deve acompanhar a proporcionalidade do impacto. Essa lógica está presente, ainda que com desenho próprio, no Regulamento Europeu de Inteligência Artificial, que distingue práticas proibidas, sistemas de alto risco, sistemas sujeitos a deveres específicos de transparência e usos de menor risco (União Europeia, 2024). Para o constitucionalismo brasileiro, a contribuição não está na importação literal desse modelo, mas na compreensão de que a intensidade dos deveres jurídicos deve variar conforme o risco de afetação de direitos fundamentais.

O risco algorítmico, enfim, não é argumento para imobilismo tecnológico; é a razão para a governança constitucional da automação. O Direito não deve impedir que sistemas automatizados auxiliem a administração pública, o mercado, o Judiciário ou os serviços essenciais, contudo, deve obstar que tais sistemas convertam pessoas em objetos de inferências opacas, estatisticamente convenientes e juridicamente injustificadas. A tarefa consiste em preservar a utilidade da técnica sem permitir que ela substitua a exigência de razões.

3 Da discriminação estrutural à discriminação algorítmica: neutralidade estatística e desigualdade material

A discriminação algorítmica só pode ser adequadamente compreendida quando se abandona a ideia de que discriminar é sempre agir com intenção subjetiva de excluir. O direito antidiscriminatório contemporâneo já demonstrou que práticas formalmente neutras podem produzir efeitos materialmente desiguais. A discriminação indireta, nesse sentido, não depende da declaração explícita de um critério proibido, mas da constatação de que determinada regra, política ou prática incide de modo desproporcional sobre grupos historicamente vulnerabilizados. A automação não elimina essa lógica, ao contrário, torna essa dinâmica mais complexa, porque distribui o tratamento desigual por meio de correlações, *proxies*, inferências e padrões estatísticos.

A discriminação estrutural é anterior ao algoritmo, isto é, forma-se em processos sociais longos, marcados por distribuição desigual de riqueza, educação, reconhecimento, moradia, segurança, trabalho, saúde e acesso institucional. Os dados coletados em sociedades desiguais não são fotografias neutras da realidade; são registros de uma realidade já atravessada por assimetrias. Quando esses dados alimentam modelos automatizados, o sistema pode aprender os padrões de exclusão que o Direito deveria combater. A promessa de neutralidade, nesse contexto, torna-se perigosa quando encobre a origem histórica dos dados.

Barocas e Selbst (2016) demonstram que o uso de *big data* pode produzir impacto discriminatório, mesmo sem intenção discriminatória explícita. Isso ocorre porque bases de dados podem ser incompletas, enviesadas ou representativas de desigualdades preexistentes; porque variáveis, aparentemente neutras, podem funcionar como aproximações de raça, de gênero, de classe, de idade ou de território; e porque a busca por maximização de eficiência pode sacrificar grupos estatisticamente minoritários ou socialmente vulneráveis. A discriminação algorítmica, portanto, não está apenas no uso de categorias sensíveis, mas também na capacidade de reconstruí-las por caminhos indiretos.

A noção de *proxy* é central. Um sistema pode não utilizar raça, gênero, religião ou origem social como variável explícita e, ainda assim, alcançar resultados discriminatórios por meio de atributos correlacionados. Endereço, histórico educacional, padrões de consumo, lacunas no currículo, tipo de dispositivo, horário de navegação, rede de contatos, trajetórias profissionais e dados comportamentais, por exemplo, podem funcionar como substitutos informacionais de categorias protegidas. Nesse contexto, a discriminação torna-se mais difícil de identificar, porque deixa de aparecer na superfície do critério e passa a operar na arquitetura da inferência.

Selbst e outros autores chamam a atenção para a dificuldade de abstrair sistemas algorítmicos de seus contextos sociais. A pretensão de corrigir tecnicamente a justiça do modelo pode fracassar quando ignora que o sistema será utilizado por instituições situadas, com objetivos próprios, incentivos econômicos, rotinas decisórias e desigualdades prévias (Selbst *et al*., 2019). Um algoritmo pode ser ajustado em laboratório para reduzir determinado viés e, ainda assim, produzir efeitos discriminatórios, quando se integra a uma organização que interpreta seus resultados de forma acrítica, escolhe determinados limiares de ação ou utiliza o modelo para legitimar decisões previamente desejadas.

Por isso, a discriminação algorítmica não deve ser definida apenas como erro do sistema. Muitas vezes, ela é o sucesso técnico de um sistema que aprendeu, fielmente, uma realidade injusta. Sob essa ótica, um modelo de recrutamento treinado sobre históricos de contratação de uma empresa que, por décadas, privilegiou homens em cargos de liderança pode reproduzir esse padrão sem qualquer ordem expressa para discriminar mulheres. Dessa forma, um sistema de concessão de crédito treinado em dados de inadimplência produzidos por um mercado historicamente desigual pode transformar vulnerabilidade econômica em risco individual naturalizado. Além disso, um sistema de policiamento preditivo, treinado sobre registros de abordagem e ocorrência, pode intensificar a vigilância em territórios mais fiscalizados, confundindo criminalidade real com presença policial pretérita.

É importante notar que a invisibilidade do fenômeno é agravada pelo vocabulário técnico. Quando uma pessoa é recusada por um modelo estatístico, o resultado pode aparecer como baixa pontuação, perfil inadequado, incompatibilidade, risco elevado, baixa aderência ou insuficiência de score. A linguagem da discriminação é substituída, então, pela linguagem da probabilidade. A consequência constitucional, porém, permanece a mesma quando a pessoa é privada de oportunidade relevante por critério que não pode ser publicamente justificado. O problema não é a probabilidade como instrumento auxiliar; é sua conversão em destino social, sem possibilidade efetiva de contestação.

Nesse ponto, a dignidade humana impede que pessoas sejam tratadas apenas como detentoras de atributos agregados. A classificação estatística pode ser útil para políticas públicas e decisões organizacionais, mas não pode substituir integralmente a consideração da pessoa concreta, quando o resultado afeta, de modo relevante, seus direitos, oportunidades ou condições de vida. Assim, a pessoa não é apenas membro provável de um *cluster*, mas é também titular de direitos, capaz de exigir razões, contestar inferências e demonstrar que a generalização não lhe é aplicável ou que o próprio critério de generalização é constitucionalmente inválido.

O direito fundamental à igualdade, por sua vez, demanda que a análise da discriminação algorítmica seja sensível a impactos. A pergunta constitucional não se limita a saber se o sistema utilizou explicitamente um critério proibido. Deve-se investigar se o resultado produziu desvantagem desproporcional, se havia alternativa menos lesiva, se os dados utilizados eram adequados, se as métricas foram estratificadas por grupo, se os afetados tinham meios de contestação, se houve avaliação prévia de impacto e se o controlador conseguiu demonstrar a legitimidade da finalidade e a necessidade do meio. A igualdade material desloca o ônus do debate: não basta afirmar que o modelo é neutro; é necessário demonstrar que seus efeitos são juridicamente justificáveis.

A discriminação algorítmica revela, assim, uma continuidade e uma ruptura. Na continuidade, as desigualdades de origem continuam influenciando o acesso a bens, a serviços e a direitos. A ruptura, por seu turno, está no modo de operação – a exclusão torna-se automatizada, escalável, opaca e revestida por autoridade técnica. Nessa transformação, o direito antidiscriminatório deve ser interpretado em diálogo com a proteção de dados, com o devido processo, com a transparência, com a proporcionalidade e com a teoria dos direitos fundamentais. Sem essa integração, corre-se o risco de enfrentar um fenômeno novo com instrumentos conceituais insuficientes.

4 Interpretação jusfundamental da discriminação algorítmica

A interpretação da discriminação algorítmica à luz dos direitos fundamentais exige uma mudança de ponto de partida. Não se deve começar pela pergunta sobre o quanto de tecnologia o Direito admite, mas pela pergunta sobre quais condições tornam legítima a utilização de tecnologia em decisões que afetam pessoas. Essa inversão é decisiva. O sistema automatizado não possui uma presunção constitucional de legitimidade apenas porque é eficiente, inovador ou estatisticamente sofisticado, visto que sua legitimidade depende da compatibilidade com a dignidade, com a igualdade, com a liberdade, com a proteção de dados, com o devido processo e com os deveres de prevenção de danos.

A dignidade da pessoa humana atua como primeiro limite material. No ambiente algorítmico, sua função não é apenas retórica, pois também impede que a pessoa seja reduzida a um perfil inferido, a uma probabilidade de risco ou a um valor econômico esperado, sem possibilidade de explicação e de contestação. Logo, a dignidade requer que a pessoa permaneça situada no centro da relação decisória, mesmo que a decisão seja mediada por sistemas técnicos. Isso significa que a automação não pode retirar do titular a possibilidade de compreender minimamente a razão da decisão, apresentar argumentos, corrigir dados, impugnar inferências e obter revisão humana quando houver impacto relevante.

A igualdade fornece o segundo eixo interpretativo. A Constituição brasileira (Brasil, [2026a]) não autoriza que diferenças socialmente produzidas sejam simplesmente convertidas em critérios privados ou estatais de classificação. Se um sistema automatizado utiliza dados que refletem desigualdades estruturais, cabe, ao agente que o emprega, demonstrar que adotou medidas razoáveis para identificar, prevenir e corrigir impactos discriminatórios. Assim, a igualdade material transforma a governança do modelo em dever jurídico, especialmente quando o sistema opera em setores que condicionam oportunidades sociais. Não se trata de exigir perfeição algorítmica, mas de impedir a indiferença institucional diante de riscos conhecidos ou previsíveis.

A proteção de dados pessoais, constitucionalizada no Brasil e densificada pela LGPD, introduz um terceiro eixo. Dados não são matéria-prima neutra; eles se vinculam à personalidade, à liberdade e à autodeterminação informativa. Nesse cenário, o tratamento automatizado de dados pessoais pode produzir efeitos intensos sobre a forma como a pessoa é vista, classificada e tratada por instituições públicas e privadas. Daí a importância dos princípios da finalidade, da adequação, da necessidade, da transparência, da prevenção, da não discriminação e da responsabilização. Esses princípios não operam como simples recomendações de boas práticas, na verdade, constituem parâmetros de validade jurídica para o tratamento automatizado (Brasil, [2026b]).

No âmbito da LGPD, o artigo 20 é particularmente relevante porque reconhece o direito de solicitar revisão de decisões tomadas unicamente com base no tratamento automatizado de dados pessoais que afetem interesses do titular, tais como decisões destinadas a definir os perfis pessoal, profissional, de consumo e de crédito ou os aspectos da personalidade (Brasil, [2026b]). Ressalte-se que esse dispositivo deve ser lido de modo constitucionalmente ampliado. A revisão não pode ser reduzida a uma confirmação formal do resultado por um humano que apenas ratifica a saída do sistema. Desse modo, revisar significa reabrir o espaço das razões, permitindo que a pessoa compreenda os critérios relevantes, conteste os dados, discuta a inferência e obtenha decisão efetivamente responsável.

O devido processo legal constitui o quarto eixo. A automação desafia o devido processo, porque leva parte da decisão para uma infraestrutura invisível. Por conseguinte, o contraditório, a ampla defesa e a motivação dependem de acesso a informações suficientes. Em matéria algorítmica, isso não significa abrir integralmente o código-fonte ou revelar todos os segredos industriais, mas disponibilizar informação suficiente para que o afetado e o órgão de controle compreendam a finalidade do sistema, as categorias de dados utilizadas, os fatores decisivos, as métricas de validação, os riscos identificados, as medidas de mitigação e os canais de revisão. Sem essa abertura mínima, não há contraditório material.

Paralelamente, a proporcionalidade oferece um método para organizar colisões. O debate sobre decisões automatizadas normalmente contempla tensões entre eficiência, inovação, segredo comercial, propriedade intelectual, segurança do sistema, proteção de dados, transparência, igualdade e liberdade econômica. A solução não pode ser binária. A proporcionalidade exige que se pergunte se o uso do sistema é adequado à finalidade legítima pretendida, se existe alternativa menos lesiva e se os benefícios justificam os custos impostos aos direitos afetados (Alexy, 2015; Barak, 2012). Quanto maior a interferência em direitos fundamentais, mais intensa deve ser a justificação.

A proporcionalidade, contudo, precisa ser complementada por uma dimensão epistêmica. Não basta ponderar valores abstratos quando as premissas técnicas são incertas, incompletas ou inacessíveis. Se o agente que utiliza o sistema não consegue demonstrar a qualidade dos dados, a validade do modelo, a estabilidade do desempenho, a inexistência de impactos discriminatórios relevantes ou a adequação das medidas de mitigação, a decisão automatizada perde suporte justificatório. Assim sendo, a incerteza técnica não pode ser transferida integralmente ao indivíduo afetado. Em situações de alto impacto, a falta de certeza deve pesar contra quem controla a infraestrutura decisória.

Essa dimensão epistêmica possibilita adaptar a lei material da ponderação aos riscos algorítmicos. Na colisão entre eficiência, inovação, segredo empresarial e proteção contra discriminação, não basta a pergunta sobre qual princípio possui maior peso em abstrato; é necessário verificar com que grau de certeza se afirmam as premissas técnicas que autorizam a restrição de direitos. A variável “C”, compreendida como grau de certeza ou confiabilidade das premissas empíricas, torna-se decisiva: quanto mais opaco, instável ou não auditado for o sistema, menor será a força justificatória da medida automatizada. Por isso, a falta de documentação, a ausência de métricas por grupo, a recusa de auditoria ou a impossibilidade de reconstrução do caminho decisório não são meros problemas administrativos, mas defeitos constitucionais de justificação (Paulo, 2026).

Tal leitura conduz a uma ideia de justificabilidade algorítmica. Uma decisão automatizada é juridicamente justificável quando pode ser reconstruída em suas razões essenciais; quando sua finalidade é legítima; quando os dados utilizados são adequados e necessários; quando os riscos foram avaliados; quando os impactos foram monitorados; quando os afetados têm meios efetivos de contestação e quando há supervisão humana capaz de revisar o resultado de modo substancial. A justificabilidade não exige que todo sistema seja simples, mas que a complexidade não elimine a responsabilidade.

A explicabilidade, nessa perspectiva, não deve ser fetichizada. Explicar não é apenas traduzir o funcionamento matemático do modelo em linguagem acessível. Em muitos casos, a explicação puramente técnica será insuficiente para o Direito. O que importa é a explicação juridicamente relevante: quais fatores influenciaram a decisão; por que tais fatores são legítimos; qual peso tiveram; quais alternativas foram consideradas; quais dados podem ser corrigidos; quais medidas foram adotadas para evitar a discriminação e como o indivíduo pode contestar o resultado. A explicabilidade constitucional é menos uma aula sobre o algoritmo e mais uma prestação de contas sobre a decisão.

Por fim, a interpretação jusfundamental da discriminação algorítmica demanda o reconhecimento da eficácia horizontal dos direitos fundamentais em contextos privados de forte poder decisório. Por exemplo, empresas de tecnologia, instituições financeiras, plataformas digitais, seguradoras, empregadores e prestadores de serviços essenciais podem exercer poderes classificatórios capazes de afetar intensamente a vida das pessoas. Quando esses poderes se apoiam em sistemas automatizados opacos, o Direito não pode tratá-los como escolhas puramente privadas. A Constituição (Brasil, [2026a]), nesse contexto, incide à medida que a decisão privada estrutura o acesso a bens fundamentais ou reproduz desigualdades materialmente relevantes.

5 Critérios de controle: explicabilidade, auditabilidade e abertura informacional proporcional ao impacto

A construção de uma resposta constitucional à discriminação algorítmica depende da definição de critérios operacionais. Sem critérios, o debate oscila entre dois extremos igualmente insatisfatórios: de um lado, a tecnofobia, que trata toda automação como ameaça; de outro, o solucionismo tecnológico, o qual presume que todo problema social pode ser resolvido por modelos melhores. O Direito precisa de um caminho intermediário, capaz de admitir usos legítimos de sistemas automatizados e, ao mesmo tempo, impor deveres robustos quando houver risco de lesão a direitos fundamentais.

O primeiro critério é a avaliação prévia de impacto. Os sistemas destinados a produzir efeitos relevantes sobre crédito, trabalho, educação, saúde, segurança pública, assistência social, migração, justiça ou acesso a serviços essenciais devem ser precedidos de análise documentada de riscos. Essa avaliação deve identificar a finalidade do sistema, os grupos potencialmente afetados, as categorias de dados utilizadas, a possibilidade de uso de *proxies* discriminatórios, os riscos de falsos positivos e falsos negativos, as alternativas menos lesivas, os mecanismos de supervisão humana e os canais de contestação. Mantelero (2018) sustenta que avaliações de impacto orientadas por direitos fundamentais são instrumentos essenciais para antecipar danos, organizar responsabilidades e evitar que a proteção seja apenas reativa.

O segundo critério é a explicabilidade suficiente. O adjetivo é importante e a explicação deve ser suficiente para a finalidade constitucional que cumpre. Para o usuário afetado, deve permitir compreender os principais fatores da decisão e os caminhos de contestação. Para a autoridade administrativa ou judicial, deve permitir controlar a legalidade do sistema, avaliar sua adequação normativa e verificar potenciais impactos discriminatórios. Para especialistas independentes, deve permitir auditoria técnica compatível com o risco. É importante destacar que a mesma informação não precisa ser entregue da mesma forma a todos, mas todos os níveis de controle devem ser funcionalmente atendidos.

O terceiro critério é a auditabilidade. Os sistemas de alto impacto devem manter documentação do ciclo de vida, registros de funcionamento, versões do modelo, bases de treinamento, critérios de validação, testes de desempenho, métricas estratificadas por grupo e histórico de alterações. Sem trilha de auditoria, a decisão automatizada torna-se evento sem memória, impossibilitando a reconstrução posterior de suas razões e falhas. A auditabilidade é, portanto, condição de responsabilização. Por meio dela, é possível distinguir erro isolado de padrão discriminatório, falha de implementação de defeito de concepção, uso indevido de problema estrutural do modelo.

O quarto critério é a supervisão humana significativa. Não basta inserir um humano ao final da cadeia decisória, se sua função for apenas confirmar o resultado do sistema. A supervisão deve ser qualificada, informada e capaz de alterar a decisão. Isso exige que o revisor humano compreenda os limites do sistema, tenha acesso às informações relevantes, disponha de autoridade institucional para discordar do resultado e seja treinado para identificar vieses, erros e impactos discriminatórios. A supervisão meramente simbólica, também conhecida como humanização de fachada, não satisfaz o devido processo algorítmico.

O quinto critério é a abertura informacional proporcional ao impacto. Em sistemas de baixo impacto, informações gerais sobre finalidade, categorias de dados e lógica básica podem ser suficientes. Em sistemas de impacto relevante, deve haver explicação individualizada, indicação de fatores decisivos e possibilidade efetiva de revisão. Em sistemas de alto impacto ou potencialmente discriminatórios, deve-se admitir auditoria independente, acesso qualificado à documentação técnica, produção de métricas por grupo e, quando necessário, inversão dinâmica do ônus informacional. A abertura não é absoluta, mas cresce conforme a gravidade da interferência.

Esse escalonamento permite harmonizar transparência e segredo comercial. A proteção da propriedade intelectual e do segredo industrial é juridicamente relevante, mas não pode funcionar como escudo para decisões que afetam direitos fundamentais sem possibilidade de controle. O próprio artigo 20, § 1º, da LGPD, ao exigir informações claras e adequadas sobre critérios e procedimentos utilizados na decisão automatizada, ressalva os segredos comercial e industrial (Brasil, [2026b]). A ressalva, contudo, não elimina o dever de informação; apenas exige que a abertura seja estruturada de modo proporcional, eventualmente por auditoria confidencial, perícia judicial, acesso restrito, relatórios técnicos ou mecanismos de proteção de sigilo.

Dando seguimento, o sexto critério é a inversão dinâmica do ônus informacional. Em litígios algorítmicos, a pessoa afetada raramente tem condições de provar a discriminação sem acesso aos elementos controlados pelo agente decisor. A exigência de prova plena do discriminado, nessa situação, equivaleria a blindar o sistema. Por isso, uma vez demonstrados indícios razoáveis de impacto desproporcional, opacidade relevante, inconsistência decisória ou negativa de informação, deve recair, sobre o controlador ou utilizador do sistema, o ônus de demonstrar a legitimidade da finalidade, a adequação dos dados, a necessidade do modelo, a inexistência de alternativa menos lesiva e as medidas adotadas para prevenir discriminação. A assimetria técnica justifica a redistribuição do ônus argumentativo e probatório.

O sétimo critério é a correção contínua. Os sistemas algorítmicos não devem ser avaliados apenas no momento de sua implantação, tendo em vista que modelos podem degradar, contextos sociais podem mudar, bases de dados podem envelhecer, padrões de uso podem se deslocar e impactos discriminatórios podem surgir

apenas após a operação em escala. A governança constitucional da automação requer monitoramento periódico, revalidação, registro de incidentes, revisão de métricas e possibilidade de suspensão do sistema, quando houver risco grave ou incerteza relevante. A prevenção não é um ato único; é um dever permanente.

A partir desses critérios, pode-se organizar um modelo de controle escalonado. O Quadro 1, a seguir, sintetiza uma proposta de gradação, sem pretensão exaustiva, destinada a orientar a intensidade dos deveres conforme o impacto potencial do sistema automatizado.

A gradação A1-A3, nesse sentido, não é uma simples classificação administrativa. Ela funciona como técnica de tradução constitucional do risco: A1 corresponde aos usos de baixa interferência, em que bastam deveres ordinários de governança; A2 corresponde a sistemas que influenciam oportunidades relevantes e demandam explicação suficiente, revisão humana e documentação técnica; A3 corresponde a sistemas de alto impacto jusfundamental, nos quais a abertura informacional reforçada, a auditoria independente, a produção de métricas estratificadas e a possibilidade de suspensão do uso tornam-se exigências diretamente associadas ao dever de proteção. O ponto decisivo do modelo consiste no fato de que risco, impacto e certeza não operam isoladamente: quanto maior o impacto e menor a certeza, mais intensas devem ser as cargas argumentativa, probatória e institucional de quem controla ou utiliza o sistema (Paulo, 2026).

Quadro 1 – Escalonamento de controle constitucional de decisões automatizadas

Grau de risco	Âmbito típico	Deveres mínimos	Resposta jurídica
A1 – risco ordinário	Automação interna, organização documental, recomendações sem efeito decisório relevante.	Informação geral, governança básica, segurança, registro de finalidade e possibilidade de correção de dados.	Controle de legalidade ordinário e responsabilização por dano demonstrado.
A2 – risco relevante	*Scores*, perfilização, triagem, ranqueamento ou recomendação com impacto sobre consumo, crédito, trabalho, educação ou serviços.	Explicação suficiente, revisão humana, documentação técnica, testes de qualidade dos dados e canal efetivo de contestação.	Redistribuição do ônus informacional diante de indícios de erro, opacidade ou impacto desproporcional.
A3 – alto risco jusfundamental	Sistemas que condicionam acesso a direitos, benefícios essenciais, saúde, segurança pública, justiça, liberdade ou grupos vulnerabilizados.	Avaliação prévia de impacto, auditoria independente, métricas estratificadas por grupo, supervisão humana significativa e abertura proporcional reforçada.	Medidas inibitórias, suspensão do uso, auditoria judicial ou administrativa e presunção desfavorável diante de resistência injustificada à informação.

Fonte: elaboração própria, adaptado de Paulo (2026).

O escalonamento não pretende substituir a análise do caso concreto. Sua função é impedir que todos os sistemas sejam tratados do mesmo modo e, sobretudo, evitar que sistemas de alto impacto recebam controle meramente formal. A categoria A1 admite governança leve, porque a interferência é reduzida; a categoria A2 exige explicação e revisão, porque o sistema já influencia oportunidades relevantes; e a categoria A3 impõe deveres reforçados, porque a decisão automatizada pode atingir o núcleo de direitos fundamentais ou grupos vulnerabilizados. O critério decisivo não é a sofisticação técnica do sistema, mas a intensidade de sua incidência sobre a vida das pessoas.

A aplicação judicial desses critérios deve evitar dois equívocos. O primeiro é transformar a explicabilidade em formalidade documental, no entanto, relatórios genéricos, termos de uso incompreensíveis ou descrições abstratas de funcionamento não bastam. A informação deve ser útil para controle e contestação. O segundo equívoco é exigir abertura total em qualquer situação, como se a proteção de segredos e da segurança do sistema não tivesse relevância jurídica. O ponto adequado está na abertura proporcional: informação suficiente para proteger direitos, com salvaguardas adequadas para preservar interesses legítimos, quando não forem incompatíveis com a tutela jusfundamental.

Esses critérios também servem à Administração Pública. O Estado não pode terceirizar sua responsabilidade decisória a sistemas automatizados contratados ou desenvolvidos internamente. Quando utiliza tecnologia para classificar cidadãos, fiscalizar condutas ou distribuir benefícios, permanece vinculado aos deveres da legalidade, da impessoalidade, moralidade, da publicidade, da eficiência, da motivação e do controle. A eficiência administrativa, embora constitucionalmente relevante, não autoriza decisões imotivadas ou discriminatórias, ao contrário, quanto maior o poder estatal de afetar direitos, maior deve ser o dever de justificar.

No âmbito privado, a mesma lógica se projeta a partir da eficácia dos direitos fundamentais e da legislação de proteção de dados. Portanto, as empresas que utilizam modelos automatizados para negar crédito, excluir

candidatos, modular preços, recusar serviços ou definir perfis de risco não podem se esconder atrás da autonomia privada quando suas decisões afetam dimensões centrais da vida social. Em suma, a livre iniciativa não se confunde com a licença para classificação opaca. Desse modo, a inovação tecnológica é constitucionalmente valiosa quando compatível com a dignidade, com a igualdade e com a proteção da personalidade.

Essa estrutura também impede que a tutela jurisdicional se limite à reparação posterior do dano. Em matéria algorítmica, a prevenção é parte do próprio conteúdo do dever de proteção: se a decisão automatizada opera em escala, com baixa visibilidade e alta assimetria informacional, o dano discriminatório tende a se disseminar antes mesmo que o indivíduo consiga compreendê-lo. A jurisdição constitucional, por isso, deve operar com medidas prospectivas e corretivas, admitindo auditorias, preservação de *logs*, exibição de documentação técnica, inversão dinâmica do ônus informacional, suspensão cautelar de sistemas de alto risco e imposição de redesenho do modelo quando a arquitetura decisória se mostrar incompatível com a igualdade material.

6 Conclusão

As decisões automatizadas recolocam um problema clássico do constitucionalismo em linguagem nova: quem decide, com quais razões, sob qual controle e com quais efeitos sobre a liberdade e a igualdade das pessoas. A automação não elimina a necessidade de justificação; apenas torna mais difícil encontrá-la, quando a decisão é conduzida para sistemas opacos, dados massivos e inferências estatísticas. É por isso que a discriminação algorítmica deve ser interpretada à luz dos direitos fundamentais, e não como simples falha técnica ou irregularidade setorial de proteção de dados.

O percurso desenvolvido permitiu afirmar que a discriminação algorítmica é uma forma contemporânea de reprodução e de intensificação de desigualdades estruturais. Sua gravidade não está apenas na possibilidade de erro, mas também na capacidade de tornar o erro escalável, invisível e institucionalmente legitimado pela aparência de neutralidade técnica. Os dados históricos podem carregar desigualdades; variáveis aparentemente neutras podem funcionar como *proxies* de categorias protegidas; modelos estatisticamente eficientes podem sacrificar grupos vulnerabilizados e decisões individualmente opacas podem produzir padrões coletivos de exclusão.

A resposta constitucional adequada exige, então, um modelo de justificabilidade algorítmica. Esse modelo parte da dignidade da pessoa humana, porque ninguém pode ser reduzido ao perfil probabilístico sem a possibilidade de compreensão e contestação; passa pela igualdade material, uma vez que impactos desproporcionais impõem uma demonstração reforçada de legitimidade; incorpora a autodeterminação informativa, pois dados pessoais estruturam a liberdade contemporânea; e se completa no devido processo, porque não há contraditório efetivo sem acesso a razões suficientes.

A transparência, nesse contexto, é necessária, mas não suficiente. O que se exige é abertura informacional proporcional ao impacto, explicabilidade juridicamente relevante, auditabilidade, supervisão humana significativa, avaliação prévia e contínua de riscos, além de redistribuição dinâmica do ônus informacional quando a prova estiver concentrada nas mãos de quem controla o sistema. Quanto maior a intensidade da interferência e menor a certeza das premissas técnicas, maior deve ser o dever de demonstração imposto ao agente público ou privado que se vale da automação.

O desafio, portanto, não é escolher entre técnica e Direito; é impedir que a técnica ocupe o lugar das razões. Os sistemas automatizados podem auxiliar decisões, organizar informações e ampliar capacidades institucionais, mas não podem transformar regularidade estatística em legitimidade normativa. A Constituição (Brasil, [2026a]) não proíbe a automação, mas proíbe que pessoas sejam governadas por classificações opacas, discriminatórias e irresponsáveis. A tarefa do Direito, nesse novo cenário, é fazer com que a inteligência das máquinas permaneça subordinada à inteligência pública dos direitos fundamentais.

Referências

ALEXY, R. **Teoria dos direitos fundamentais**. Tradução de Virgílio Afonso da Silva. 2. ed. São Paulo: Malheiros, 2015.

BARAK, A. **Proportionality:** constitutional rights and their limitations. Cambridge: Cambridge University Press, 2012. DOI: https://doi.org/10.1017/CBO9781139035293

BAROCAS, S.; SELBST, A. D. Big data's disparate impact. **California Law Review**, Berkeley, v. 104, n. 3, p. 671-732, 2016. DOI: https://doi.org/10.2139/ssrn.2477899

BRASIL. [Constituição (1988)]. **Constituição da República Federativa do Brasil**. Brasília, DF: Presidência da República, [2026a]. Disponível em: https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 21 jun. 2026.

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF: Presidência da República, [2026b]. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 21 jun. 2026.

BURRELL, J. How the machine 'thinks': understanding opacity in machine learning algorithms. **Big Data & Society**, London, v. 3, n. 1, p. 1-12, 2016. DOI: https://doi.org/10.1177/2053951715622512

MANTELERO, A. AI and Big Data: a blueprint for a human rights, social and ethical impact assessment. **Computer Law & Security Review**, Amsterdam, v. 34, n. 4, p. 754-772, 2018. DOI: https://doi.org/10.1016/j.clsr.2018.05.017

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

PAULO, L. M. **Discriminação algorítmica e proteção de grupos em situação de vulnerabilidade pela jurisdição constitucional:** dever de proteção estatal à luz da lei material da ponderação. 2026. Tese (Doutorado em Direito) – Programa de Pós-Graduação em Direito, Universidade de Santa Cruz do Sul, Santa Cruz do Sul, 2026. Disponível em: https://repositorio.unisc.br/jspui/handle/11624/4267. Acesso em: 21 jun. 2026.

SELBST, A. D.; BOYD, D.; FRIEDLER, S. A.; VENKATASUBRAMANIAN, S.; VERTESI, J. Fairness and abstraction in sociotechnical systems. *In*: CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY, 2019, Atlanta. **Proceedings** [...]. New York: ACM, 2019. p. 59-68. DOI: https://doi.org/10.1145/3287560.3287598

UNIÃO EUROPEIA. Regulamento (UE) 2016/679 do Parlamento Europeu e do Conselho, de 27 de abril de 2016, relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados e que revoga a Diretiva 95/46/CE (Regulamento Geral sobre a Proteção de Dados). **Jornal Oficial da União Europeia**, Bruxelas, série L, n. 119, p. 1-88, 4 maio 2016. Disponível em: http://data.europa.eu/eli/reg/2016/679/oj. Acesso em: 21 jun. 2026.

UNIÃO EUROPEIA. Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho, de 13 de junho de 2024, que estabelece regras harmonizadas em matéria de inteligência artificial e que altera os Regulamentos (CE) n.o 300/2008, (UE) n.o 167/2013, (UE) n.o 168/2013, (UE) 2018/858, (UE) 2018/1139 e (UE) 2019/2144 e as Diretivas 2014/90/UE, (UE) 2016/797 e (UE) 2020/1828 (Regulamento da Inteligência Artificial). **Jornal Oficial da União Europeia**, Bruxelas, série L, p. 1-144, 12 jul. 2024. Disponível em: http://data.europa.eu/eli/reg/2024/1689/oj. Acesso em: 21 jun. 2026.