

Responsabilidade civil e riscos discriminatórios na tomada de decisão pela inteligência artificial: desafios e perspectivas no direito brasileiro

Civil Liability and Discrimination Risks in Artificial Intelligence Decision-Making: Challenges and Prospects in Brazilian Law

Crys Frydolain Banangouna¹ e João Dantas Pereira²

v. 14/ n. 3 (2026)
Julho/Setembro

Aceito para publicação em 16/06/2026.

¹Graduando em Direito pela Universidade Federal do Rio Grande do Norte, Natal, Rio Grande do Norte. ORCID: 0009-0002-0578-5540. E-mail: banangounacrysfrydolain@gmail.com;

²Doutorado em Sociologia pela Universidade de São Paulo (1999). Mestrado em Sociologia pela Universidade de São Paulo (1989). Graduação em Direito pela Faculdade Estácio de Natal (2014). Graduação em Serviço Social pela Universidade Federal do Rio Grande do Norte (1981). Atualmente é Coordenador do Curso de Relações Internacionais. Professor Associado, no Departamento de Direito Processual e Propedêutica/Universidade Federal do Rio Grande do Norte. Tem experiência na área de Direito, com ênfase em Filosofia do Direito e Ética Geral e Profissional. Desenvolveu/desenvolve estudos nos seguintes temas: alcoolismo, violência, Relações Internacionais. Foi avaliador de cursos de Direito e Serviço Social/Instituto Nacional de Ensino e Pesquisa/Ministério da Educação e Cultural/Brasil. ORCID: 0000-0001-9855-827X. E-mail: jotarn08@gmail.com.

<httpswww.gvaa.com.br/revista/index.php/RDGP>

RESUMO: O avanço dos sistemas de Inteligência Artificial (IA) em domínios sensíveis (recrutamento, crédito, segurança pública e decisões judiciais) impõe desafios à responsabilidade civil por danos discriminatórios. Este artigo objetiva verificar se o ordenamento jurídico brasileiro Código Civil (arts. 186 e 927), Lei Geral de Proteção de Dados (LGPD) e Código de Defesa do Consumidor (CDC) tutela adequadamente vítimas de discriminação algorítmica. A pesquisa, de metodologia dedutiva e abordagem qualitativa, fundamenta-se em revisão bibliográfica nacional e estrangeira, análise comparada da legislação e exemplos de casos concretos (Amazon e Meta/ANPD). Os resultados mostram que o viés algorítmico, enraizado em dados históricos, atua como risco sistêmico, reproduzindo preconceitos de raça, gênero e classe. A teoria subjetiva (art. 186, CC) mostra-se ineficaz devido à opacidade dos modelos e à insuperável dificuldade probatória do nexo causal pela vítima. Embora a responsabilidade objetiva pelo risco (art. 927, CC), a inversão do ônus da prova (arts. 42, LGPD; 6º, CDC) e o direito de revisão (art. 20, LGPD) representem avanços, sua aplicação isolada esbarra em limitações como a ausência de auditorias compulsórias e a fragilidade da ANPD. O PL 2.338/2023, inspirado no AI Act europeu, avança na classificação de risco, mas silencia sobre danos existenciais difusos.

Conclui-se que o sistema de vigente é insuficiente para garantir a reparação integral. A tutela efetiva exige três medidas: fortalecimento da ANPD para auditorias obrigatórias; previsão de seguro civil no PL 2.338/2023; e enunciado jurisprudencial fixando a inversão do ônus da prova como regra geral.

Palavras-chave: Discriminação algorítmica. Responsabilidade civil. Inteligência artificial. LGPD. Viés algorítmico. Direito digital.

ABSTRACT: The advancement of Artificial Intelligence (AI) systems in sensitive domains (recruitment, credit, public security, and judicial decisions) poses challenges to civil liability for discriminatory damages. This article aims to verify whether the Brazilian legal framework – namely the Civil Code (arts. 186 and 927), the General Data Protection Law (LGPD), and the Consumer Protection Code (CDC) – adequately protects victims of algorithmic discrimination. The research adopts a deductive methodology with a qualitative approach, grounded in a review of national and international literature, comparative legislative analysis, and case studies (Amazon and Meta/ANPD). The results indicate that algorithmic bias, rooted in historical data, functions as a systemic risk, reproducing prejudices of race, gender, and class. The subjective theory of liability (art. 186, CC) proves ineffective due to model opacity and the insurmountable difficulty for victims to prove the causal link. Although strict liability for risk (art. 927, CC), the reversal of the burden of proof (arts. 42, LGPD; 6, CDC), and the right to review (art. 20, LGPD) represent progress, their isolated application faces limitations such as the lack of mandatory audits and the institutional fragility of the ANPD. Bill 2.338/2023, inspired by the European AI Act, advances risk classification but remains silent on diffuse existential damages. It is concluded that the current system is insufficient to guarantee full compensation. Effective protection requires three measures: strengthening the ANPD to carry out

mandatory audits; including compulsory civil insurance in Bill 2.338/2023; and issuing a judicial precedent establishing the reversal of the burden of proof as a general rule.

Keywords: Algorithmic discrimination. Civil liability. Artificial intelligence. LGPD (Brazilian General Data Protection Law). Algorithmic bias. Digital law.

1. CONSIDERAÇÕES INICIAIS

A Inteligência Artificial (IA) tem se tornado uma presença ubíqua na sociedade contemporânea, assumindo papel central tanto nas atividades cotidianas quanto em setores estratégicos da economia e da administração pública. Desde suas primeiras aplicações até os atuais modelos generativos e de aprendizado profundo, a tecnologia evoluiu para ferramentas capazes de realizar tarefas complexas que, em muitos casos, superam nossa capacidade cognitiva. Essa evolução, embora acompanhada de inegáveis benefícios como a análise de grandes volumes de dados e a otimização de serviços, também traz consigo desafios jurídicos. Com o desenvolvimento de técnicas de machine learning e deep learning, os sistemas de IA passaram a operar com elevado grau de autonomia, aprendendo com experiências e tomando decisões com reduzida ou nenhuma interferência humana. Atualmente, tais sistemas são amplamente utilizados tanto no setor privado quanto no público, em atividades como seleção de candidatos a emprego, análise de perfil de crédito, concessão de benefícios sociais, apoio à decisão judicial e vigilância por reconhecimento facial. A eficiência dessas ferramentas, contudo, não as isenta de um problema estrutural: ao serem treinados com dados históricos que refletem desigualdades sociais, culturais e organizacionais, os algoritmos tendem a absorver e reproduzir os preconceitos da sociedade que os gerou.

A literatura especializada documenta amplamente esse fenômeno a automatização da discriminação em que sistemas reproduzem e até amplificam vieses de raça, gênero, classe social e origem, sem que qualquer agente humana tenha programado deliberadamente para esse fim. No Brasil, essa realidade assume contornos ainda mais graves, em razão das profundas desigualdades sociais e econômicas que marcam o país. Como alerta Silvio Almeida (2019), racismo não se limita a comportamentos individuais, mas é resultado do funcionamento estrutural das instituições econômicas, políticas e jurídicas que o reproduzem como componente orgânico da ordem social. Essa premissa estrutural projeta-se diretamente sobre os sistemas de IA: quando treinados com dados produzidos por essas mesmas instituições históricas, os algoritmos tendem a replicar e amplificar as hierarquias neles inscrita conclusão que encontra confirmação empírica na literatura especializada sobre discriminação algorítmica (Wang et al., 2024; Buolamwini & Gebru, 2018). Essa tese encontra confirmação empírica no estudo "Gender Shades", de Joy Buolamwini e Timnit Gebru (2018), que evidenciou taxas de erro de até 34,7% em sistemas comerciais (IBM; Microsoft; Face; Google) de classificação de

gênero por análise facial para mulheres de pele mais escura, contra 0,8% para homens de pele mais clara demonstrando que disparidades de desempenho se acumulam interseccionalmente sobre grupos historicamente marginalizados (Buolamwini; Gebru, 2018). Na mesma direção, Bianca Kremer Nogueira Correa (2021) problematiza a posição das populações afrolatinas diante dessas tecnologias, mostrando que a discriminação algorítmica assume contornos específicos no contexto das Américas

Diante desse cenário, surge a pergunta central que orienta este trabalho: o atual sistema brasileiro de responsabilidade civil consagrado pelo Código Civil (CC), pela Lei Geral de Proteção de Dados (LGPD) e pelo Código de Defesa do Consumidor (CDC) consegue proteger adequadamente as vítimas de discriminação algorítmica? Nossa hipótese é o atual sistema de responsabilidade civil, embora formalmente apto, mostra-se materialmente insuficiente para tutelar as vítimas de discriminação algorítmica.

Do ponto de vista metodológico, adota-se pesquisa qualitativa, de natureza dedutivo, com as seguintes etapas: a) revisão bibliográfica e documental em doutrina nacional e estrangeira, com predomínio de fontes publicadas e selecionadas por pertinência temática (discriminação algorítmica, responsabilidade civil, proteção de dados e governança da IA); b) análise de legislação comparada, com foco no Regulamento 2024/1689, no Projeto de Lei brasileiro nº 2.338/2023 (em tramitação) e no Projeto de Lei nº 4/2025 (reforma do Código Civil); c) identificação e tratamento crítico de quatro casos concretos Amazon, Gender Shades, sistema ASSIS do Tribunal de Justiça do Estado do Rio de Janeiro e atuação da ANPD no caso Meta(selecionados em razão de sua relevância paradigmática para o debate brasileiro e por sua diversidade setorial (trabalho, reconhecimento facial, jurisdição e proteção de dados). O desenho da pesquisa segue a tradição metodológica de estudos jurídicos propositivos (Estêves; João, 2025; Frazão, 2024) e busca não apenas descrever o problema, mas também oferecer parâmetros operativos para a reforma em curso.

A estrutura do trabalho reflete a própria lógica da argumentação além desta introdução e da conclusão: o Capítulo 2 examina os mecanismos de discriminação algorítmica; o Capítulo 3 analisa as insuficiências do modelo tradicional de responsabilidade civil; o Capítulo 4 avalia as soluções existentes no arcabouço legal brasileiro; o Capítulo 5 discute os projetos de lei em tramitação; e o Capítulo 6 investiga as implicações práticas nos setores trabalhista, consumerista e administrativo.

2. A DISCRIMINAÇÃO COMO RISCO SISTÊMICO DOS MODELOS AUTÔNOMOS

2.1. Mecanismos de aprendizado enviesado

Antes de falarmos sobre discriminação algorítmica como um problema legal, precisamos entendê-la como um problema que surge não apenas de uma questão técnica, mas sim de uma falha técnica e de um problema com a estrutura de dados e a sociedade que a produziu. Quando sistemas de IA são treinados para tomar decisões sobre classificação de risco de crédito, seleção de candidatos, priorização de pacientes ou avaliação de desempenho, eles aprendem padrões estatísticos a partir de dados gerados por instituições historicamente excluídas do sistema social ou econômico.

Como bem estabelecido pela doutrina, modelos treinados em dados históricos, que muitas vezes são baseados em dados, tendem a fortalecer e reforçar preconceitos sociais e não a eliminá-los (Wang et al., 2024). Pesquisas especializadas identificaram três tipos de viés na literatura: viés de dados (erro na fonte de dados), viés de design (processos realizados no design do modelo) e ciclo de feedback (ciclo vicioso, onde a saída é retroalimentada nas entradas subsequentes; Wang et al., 2024; Mendes; Mattiuzzo, 2019) são todos comuns e trabalham juntos em cada sistema.

Selena Silva e Martín Kenney, em seu estudo sobre teoria econométrica e institucional em plataformas digitais, demonstram que o viés étnico não está localizado apenas nos dados em um ponto do processo, mas é consequência do impacto cumulativo dos dados históricos, da prática de engenharia e da ausência de governança no sistema (Selena; Kenney, 2018). Entretanto, tais procedimentos induzem tais sistemas a dar mais importância para padrões probabilísticos fixos em vez das características únicas de cada cenário. Em consequência, quando os dados integrados ao sistema veiculam estereótipos ligados a um grupo social ou a um gênero, a ferramenta está propensa a repetir e fazer crescer esses desequilíbrios, podendo, assim, gerar resultados de caráter discriminatório.

Esta lógica estatística do aprendizado de máquina gera médias normativas nas quais o indivíduo é julgado não por suas próprias características, mas pela probabilidade estimada de pertencer ao grupo. Isso é o que se refere como "injustiça por generalização" (Doneda et al., 2018). Essa vulnerabilidade teórica encontra confirmação empírica no estudo publicado em 2019 na revista "Science" por cientistas americanos, que evidenciou um viés racial crítico em um algoritmo comercial amplamente utilizado para priorizar pacientes com necessidades complexas de saúde. Os pesquisadores demonstraram que, para uma mesma pontuação de risco gerada pelo software, os pacientes negros estavam significativamente mais doentes do que os pacientes brancos, uma disparidade que reduzia consideravelmente o acesso do grupo negro a programas de cuidados adicionais. O erro fundamental do sistema, como destaca o estudo, residia na escolha de "proxies » (variáveis substitutas) convenientes: ao prever os custos futuros com saúde em vez das necessidades médicas reais, o algoritmo

ignorou o fato de que as desigualdades estruturais de acesso levam o sistema a gastar historicamente menos com pacientes negros, tratando o menor uso de recursos como um falso indicador de melhor saúde. (Obermeyer et al., 2019).

Esta é a discriminação algorítmica não necessariamente baseada em uma intenção consciente de discriminar, mas nas próprias regras dos sistemas de IA, e a discriminação pode surgir da lógica desse aprendizado estatístico, não importa quão malicioso seja. Nessa perspectiva, Laura Mendes e Marcella Mattiuzzo (2019) sistematizam o fenômeno em:

quatro categorias fundamentais que evidenciam o funcionamento dessa lógica: a discriminação por erro estatístico, decorrente de falhas nos dados ou nos códigos; a discriminação por generalização, que prejudica indivíduos atípicos incluídos em médias estatisticamente corretas; a discriminação por uso de informações sensíveis ou protegidas (proxies); e a discriminação que limita o exercício de direitos, que penaliza o titular pelo simples fato de exercer suas prerrogativas legais. (Mendes; Mattiuzzo, 2019. p 51-53)

A classificação de Mendes e Mattiuzzo embora didática, não se estende aos diferentes tipos de viés algorítmico. Como argumenta Gabrielle Johnson na obra de referência, o viés algorítmico não é um "preconceito da máquina", mas sim um viés tecnológico social por si só, surgindo de um processo de informação muito inocente, não da intenção dos programadores, e é muito difícil de detectar e corrigir (Johnson, 2021).

É importante, ressaltar que mostrar sinais estatísticos confiáveis em um conjunto de dados com milhões de observações, não é uma tarefa fácil, e na prática, modelos com milhões de parâmetros são difíceis de interpretar. O primeiro é a substituição de categorias protegidas por proxies aparentemente neutros: códigos postais, latitude ou camada de renda como proxies para raça em modelos de pontuação de crédito é o exemplo mais conhecido na literatura nacional (Ferreira & Prado, 2024; Vilarino & Vicente, 2020). O segundo é a generalização sistemática de perfis médios em grandes grupos demográficos que o sistema nunca pode discriminar internamente (como inadimplência racializada): a entrada do modelo é a inadimplência média histórica de uma comunidade, não um nível específico de pagamento (Ferreira & Prado, 2024). O terceiro emergente é o reforço institucional: decisões automatizadas em segurança pública, seleção de pessoal e concessão de crédito, se aplicadas em escala e alimentando novos bancos de dados, acabam por produzir exatamente o que esperavam.

A literatura especializada também mostra que as próprias métricas de justiça algorítmica usadas em auditorias técnicas o cerne do debate europeu, mas cada vez mais integradas na doutrina brasileira podem manter vieses estruturais (Wachter et al., 2021). Wachter, Mittelstadt e Russell, ao analisarem a compatibilidade legal dessas duas famílias de métricas de justiça e a lei europeia de antidiscriminação, descobriram que essas métricas não abordam o viés estrutural já existente, pois

não visam a igualdade e apenas a inclusão delas não é suficiente por si só e não é uma maneira fácil de determinar a conformidade (Wachter et al., 2021).

2.2. Exemplos concretos: nacionais e internacionais

Para ilustrar a operacionalização do viés algorítmico na prática, examinam-se a seguir casos emblemáticos ocorridos em diferentes setores. No plano internacional, o caso Amazon tornou-se referência obrigatória: a empresa desenvolveu um sistema de triagem de currículos que penalizava documentos contendo a palavra “mulheres” (como em “capitã do time de mulheres”) e desvalorizava candidatas de universidades femininas, porque o modelo havia sido treinado com dez anos de contratações históricas majoritariamente masculinas. A empresa descontinuou o projeto, mas o caso ilustra como o viés nos dados de origem se combina com problemas no design do sistema (Mateus, 2018).

No contexto brasileiro, o sistema de Apoio ao Sistema de Segurança e Inteligência Sistêmica (ASSIS), desenvolvido pelo Tribunal de Justiça do Rio de Janeiro para apoiar decisões em ações de reintegração de posse, foi alvo de críticas por potencialmente privilegiar argumentos de proprietários em detrimento de comunidades vulneráveis, revelando como o viés se infiltra até mesmo nos sistemas preditivos utilizados no Poder Judiciário (Santiago, 2025). Em credit scoring, setor de enorme relevância para a mobilidade econômica (Ferreira; Prado 2024), estudos mostram que esses algoritmos de crédito acabam atribuindo pontuações menores a moradores de certas regiões geográficas, reproduzindo padrões históricos de exclusão financeira que, no Brasil, prejudicam muito mais a população negra e da periferia.

Esses exemplos demonstram que, do ponto de vista teórico, a discriminação algorítmica não se encaixa no modelo canônico do ilícito civil, baseado em ação humana consciente e intencional. Trata-se de um dano sistêmico e difuso, que muitas vezes só aparece nas estatísticas, afetando categorias de pessoas de forma probabilística, sem que qualquer decisão individual seja, em si, identificável como ilícita. Essa característica é central para compreender os desafios que o próximo capítulo vai abordar.

2.3. Violação de direitos fundamentais e a Constituição Federal de 1988

E se lermos a Constituição Federal de 1988 sob a perspectiva da discriminação algorítmica, vemos que o sistema constitucional brasileiro não é indiferente ao tema, há cláusulas explícitas no texto que são contra sistemas baseados em hierarquias sociais. O artigo 5º, por exemplo, garante a igualdade e proíbe distinções baseadas em raça, sexo, cor, origem ou qualquer outra condição pessoal, e isso elimina a indiferença legal à discriminação produzida por sistemas algorítmicos baseados nessas mesmas características, mesmo que não sejam projetados para um propósito estatístico. Por outro lado, o artigo 5º, X, para que os dados pessoais sejam mantidos na vida privada, de acordo com a

doutrina, garante proteção constitucional contra a coleta sistemática de dados, tendo em vista os debates mais complexos na doutrina brasileira sobre big data no governo (Araújo et al., 2020). O artigo 5º, LXXIX, através da Emenda Constitucional nº 115/2022, colocou os dados pessoais como um direito fundamental autônomo, e seu uso em perfis algorítmicos foi importante e justificado.

Outra cláusula relevante, ainda debatida em debates doutrinários mais recentes: o princípio articulado no Artigo 3º, IV, da Constituição Federal de 1988, um objetivo da República que se funda nos preceitos: “promover o bem-estar de todos, sem preconceitos de origem, raça, sexo, cor, idade e quaisquer outras formas de discriminação”. Essa norma programática constitui uma força vinculante legal e direciona tanto o Estado quanto os indivíduos a um dever positivo de igualdade material. O design, implementação ou supervisão de sistemas automatizados deve atender a essa obrigação, caso contrário, sua ausência por si só fornece fundamentos para responsabilidade civil.

Luis Roberto Barroso e a doutrina da constitucionalização do Direito Civil sugerem que, juntos, esses dispositivos produzem a função diretamente constitucional de disposições que previnem a discriminação, mesmo sem a existência de legislação infraconstitucional específica (Rodrigues & Leal, 2018). Enquanto Rodrigues e Leal, em uma exploração mais aprofundada da jurisprudência do STF sobre o efeito horizontal dos direitos fundamentais, afirmam que, com o impacto de uma decisão histórica no RE 201.819-8, abriu-se a porta para a responsabilização e responsabilidade de empresas privadas por violações de direitos que são explicitamente proibidas pela Constituição (argumentam os autores Rodrigues e Leal, 2018), isso também se alinha com a tese do Supremo Tribunal Federal de que os direitos constitucionais são inseparáveis das relações entre indivíduos, mas que essas relações são particularmente fortalecidas quando existem assimetrias factuais estruturais, como ocorre quando a interação está ocorrendo entre um indivíduo e um operador de algum sistema automatizado de tomada de decisão.

Silvério Luiz de Almeida, por sua vez, ao discutir o racismo no Brasil, observa que ele é uma parte estrutural das instituições econômicas, políticas e jurídicas e não apenas sobre o comportamento individual. A relação entre discriminação algorítmica e a Constituição de 1988 não é apenas acidental na aplicação de cláusulas de proteção, mas requer uma leitura sistêmica voltada para o desmantelamento institucional do racismo estrutural (Almeida, 2019). Para o autor, a questão constitucional não é apenas evitar a discriminação, mas também desmantelar os mecanismos institucionais que a alimentam, e essa é a tarefa digital do Estado regulador e dos operadores privados.

3. O SISTEMA DE RESPONSABILIDADE CIVIL EM XEQUE

3.1. Inadequação da responsabilidade subjetiva

A teoria subjetiva implica que, para se realizar um ato ilícito civil, é necessário reunir quatro elementos: conduta, culpa, nexo causal e danos. Nos sistemas de tomada de decisão algorítmica, no entanto, três desses elementos são praticamente inacessíveis para a vítima ou para o intérprete legal em termos de provas materiais (Vieira & Siqueira, 2023; Zech, 2021).

Essas três barreiras transformam o instituto clássico em um mecanismo formalmente disponível, mas materialmente inadequado (Buiten, 2024). No entanto, essa estrutura normativa não foi projetada para lidar com danos causados por sistemas autônomos de Inteligência Artificial.

A literatura americana sobre este assunto foi desenvolvida a partir do campo da inteligência artificial clínica, onde a opacidade do raciocínio automatizado e a dificuldade de prova são paralelas aos problemas de discriminação algorítmica geral. Quando Sullivan e Schweikart (2019) analisam a inadequação dos regimes tradicionais de responsabilidade civil para sistemas clínicos automatizados, eles constataam que a opacidade e a autonomia desses sistemas tornam virtualmente impossível para a vítima individual provar a culpa (Sullivan; Schweikart, 2019). O diagnóstico mais recente nessa direção foi desenvolvido na responsabilidade médica por IA, mas o argumento estrutural desses autores, de que os regimes baseados em culpa se fundamentam em um acesso epistêmico ao processo de tomada de decisão que é negado pelos sistemas algorítmicos é portanto, aplicável ao problema da discriminação algorítmica (crédito, emprego e segurança pública) e as barreiras à prova são tanto estruturais quanto acidentais e que as mesmas dificuldades se repetem no campo da discriminação algorítmica, já que os sistemas autônomos de aprendizado de máquina não são bem definidos (Vieira; Siqueira, 2023; Zech, 2021).

Nesse sentido, Oscar Valente Cardoso (2025) sintetiza o problema nos seguintes termos:

A crescente integração da inteligência artificial em diversos setores da sociedade traz consigo um conjunto complexo de desafios regulatórios que exigem atenção por parte de formuladores de políticas públicas [...] A natureza disruptiva da inteligência artificial, aliada à sua rápida evolução, torna difícil a tarefa de criar regulamentos que sejam ao mesmo tempo eficazes e flexíveis o suficiente para se adaptarem a novas realidades. Além disso, a IA levanta questões específicas sobre responsabilidade, transparência e ética, que não são facilmente resolvidas com fundamento nos marcos regulatórios tradicionais (Cardoso, 2025, p.56)

Com este diagnóstico, parte da doutrina defende a modernização do Código Civil e que a discriminação probatória pode ser mitigada se a jurisprudência integrar deveres específicos de cuidado para a Inteligência Artificial (por exemplo, auditoria prévia antes do lançamento, monitoramento contínuo, boa supervisão humana e rastreabilidade de decisões) (Doneda et al., 2018). Sob este ponto de vista, a responsabilidade desloca-se do resultado discriminatório para o comportamento do

desenvolvedor, e a falha não seria a entrega de um resultado específico, mas sim a falta de governança, supervisão e curadoria de dados para prevenir vieses.

Essa abordagem assume que é possível encontrar no fluxo do sistema o momento exato e que o agente pode ser responsável por qualquer possível falha no dever de cuidado, sendo que a opacidade algorítmica e a autonomia dos modelos torna isso virtualmente impossível. Na verdade, a desvantagem dessa abordagem é que as vítimas precisam provar não apenas o abuso, mas também o dever de cuidado que o agente não conseguiu manter, e em que ponto do ciclo de vida do sistema. Para uma vítima que é discriminada em um contexto algorítmico, que provavelmente será uma vítima socialmente vulnerável com recursos legais e técnicos limitados à sua disposição, há uma grande dificuldade a ser resolvida na prática (Rodrigues, 2020).

3.2. O problema probatório e a opacidade algorítmica

A opacidade dos sistemas de Inteligência Artificial, em particular dos modelos de aprendizado profundo, é o principal obstáculo técnico à responsabilidade civil. Esses sistemas, construídos com base em uma combinação de dados diversos, melhorias técnicas e parâmetros complexos cuja natureza e base legal muitas vezes estão ocultas e não são explicadas de forma inteligível, nem sequer fornecem uma explicação coerente de cada uma das decisões individuais que tomam, o que a doutrina chama de “caixa-preta”: uma opacidade algorítmica que torna impossível rastrear a cadeia causal que gerou o resultado (Vieira; Siqueira, 2023).

Dessa forma, do ponto de vista prático, a vítima de uma decisão algorítmica discriminatória, como um candidato a emprego rejeitado por um sistema de triagem automatizado, não pode determinar se a decisão foi resultado de uma falha técnica do modelo, de uma adequação ao perfil do trabalho ou, na realidade, de um critério discriminatório oculto. Luciano Floridi (2023) propõe a explicabilidade como o quinto princípio central de sua estrutura ética de IA, juntamente com beneficência, não maleficência, autonomia e justiça, e argumenta que a inteligibilidade nos processos algorítmicos é necessária para o uso responsável da IA e para a responsabilização dos algoritmos envolvidos (Floridi, 2023). Este princípio ético também é um pré-requisito para a transparência processual: se a decisão automatizada não for explicada adequadamente, a pessoa que está contestando a decisão não consegue contestá-la e, portanto, perde o direito de revisão da LGPD.

Duas soluções doutrinárias foram propostas para aliviar essa dificuldade probatória. A primeira é a inversão do ônus da prova: em vez de ser necessário mostrar o viés do algoritmo, o desenvolvedor ou usuário deve provar que o sistema é justo, auditável e não discriminatório (Gerards;

Xenidis, 2021). Esta disposição já é apoiada pelo Artigo 42, § 2, da LGPD, que permite a inversão do ônus da prova em favor do titular dos dados em uma investigação sobre o processamento de dados pessoais. A segunda é a teoria dos ônus dinâmicos da prova, que atribui o ônus probatório de acordo com a posição epistêmica das partes, ou seja, quem tem a capacidade técnica e acesso ao código e dados (o provedor do sistema) deve demonstrar a regularidade e confiabilidade do algoritmo. Ambas são apoiadas no Artigo 373, § 1, do Código de Processo Civil e permitem que o juiz distribua o ônus da prova de forma diferente, dependendo da facilidade ou dificuldade de provar a evidência.

3.3. Breve menção à (des) necessidade de personalidade jurídica da IA

Outra questão que merece breve menção, embora não seja central para os fins deste trabalho é a possibilidade de atribuir personalidade jurídica aos sistemas de IA. No que diz respeito à doutrina brasileira atual e ao direito comparado (Projeto de Lei 2.338/2023 e Lei de IA da UE), essa abordagem não é aplicável para responsabilidade civil (Barbosa, 2025). Isso ocorre porque dar personalidade à IA é o mesmo que proteger o desenvolvedor e os usuários de uma entidade cujos bens não são próprios, e na prática tornaria impossível para a vítima receber indenização pelo dano. Portanto, a visão atual é que os sistemas de IA devem ser tratados como ferramentas tecnologicamente avançadas e a responsabilidade de arcar com a responsabilidade legal recairá sobre pessoas físicas ou jurídicas (Henz, 2021).

4. SOLUÇÕES EXISTENTES NO ORDENAMENTO BRASILEIRO E SUAS LIMITAÇÕES

4.1. Responsabilidade objetiva pelo risco (art. 927, parágrafo único, CC)

A teoria objetiva da responsabilidade, prevista no artigo 927, parágrafo único do Código Civil, dispensa a demonstração de culpa pelo agente quando a atividade que ele ou ela realiza é perigosa para os direitos de terceiros. A aplicação desta cláusula geral de risco ao desenvolvimento e uso de sistemas de Inteligência Artificial, torna-se um caminho mais promissor no quadro normativo existentes. O debate atual sobre a regulação de riscos é fortemente influenciado pelo AI Act da União Europeia e pelo Projeto de Lei nº 2.338/2023 em tramitação, que afirmam que a responsabilidade deve ser graduada com base no nível de risco e autonomia do sistema. Assim, sistemas de alto risco, como recrutamento, concessão de crédito, segurança pública e decisões judiciais, estariam sujeitos a um regime de responsabilidade objetiva plena (Frazão, 2024), enquanto sistemas de baixo risco com potencial menos nocivo poderiam estar sujeitos a um regime menos rigoroso (Viana; Toscano, 2025).

Embora a classificação de risco seja discutida como proporcionalidade e a inovação seja incentivada, os requisitos regulatórios tendem a ser mais rigorosos para sistemas que podem causar danos significativos. O verdadeiro desafio, consta, em primeiro lugar na identificação do que constitui sistemas de alto risco de acordo com as atividades realizadas. Existem tantas aplicações de IA e

diferentes tipos de aplicações, como sistemas de recomendação comercial a sistemas de tomada de decisão na área da saúde, que não é razoável estabelecer limites rígidos e rápidos para atividades de risco (Frazão, 2024). Enquanto o Artigo 927, o parágrafo único do Código Civil funciona bem para atividades economicamente definidas, as tecnologias transversais precisam de construção jurisprudencial (Zech, 2021).

Em segundo lugar, na persistência do problema probatório. Mesmo sob o regime objetivo, ainda é exigido da vítima provar o dano, onexo causal, e em menor grau, o comportamento do agente. Em terceiro lugar, na fragilidade do sistema de auditoria corporativa, que é, no regime privado brasileiro, a principal ferramenta de controle de risco. Muito está sendo feito em contratos médios e grandes, mas a auditoria algorítmica não é bem regulamentada no Brasil. De fato, a lacuna regulatória foi parcialmente preenchida pela ANPD recentemente, especialmente no caso Meta/2024, onde a agência suspendeu preventivamente a política de privacidade devido ao risco concreto e elevado de dano (Camargo, 2024; Souza et al., 2024).

Nesta situação, a teoria do risco não é suficiente. Sua aplicação à inteligência artificial requer complemento normativo em três áreas: definição legislativa dos critérios de risco material, seguindo o PL 2.338/2023 (Viana; Toscano, 2026); instrumentos processuais específicos para demonstração clara de opacidade e nexocausal (Estêves; João, 2025; Mulholland, 2025); e uma estrutura institucional que possa produzir auditorias quantitativas em larga escala (Camargo, 2024; Sarlet; Rodrigues, 2022).

No modelo objetivista, há também responsabilidade solidária entre o desenvolvedor e o usuário. Assim, o desenvolvedor que cria um sistema de Inteligência Artificial é responsável pelos defeitos originais e o usuário por quaisquer falhas que foram criadas no processo de adaptação (Buiten, 2024). Essa responsabilidade solidária, na forma dos Artigos 264 a 285 do Código Civil e na forma do parágrafo único do Código de Defesa do Consumidor, permite que a vítima processe qualquer uma das partes responsáveis pelo defeito, não importando quando o defeito ocorreu na cadeia. No entanto, seu uso é desafiado quando o sistema com aprendizado autônomo é variável e o resultado não é a configuração original, sendo, portanto, difícil determinar a responsabilidade entre os diferentes agentes da cadeia de produção (Affonso, 2025).

4.2. LGPD: arts. 20, 42 e 6, IX (escopo e limites)

O artigo 20 da LGPD concede o direito de revisar decisões automatizadas. Isto é, o reconhecimento importante na legislação brasileira do que a literatura europeia chama de “direito à explicação”: o direito de ter uma explicação clara e adequada da lógica por trás da decisão de automatizar

completamente (Dourado e Aith, 2022; Laux, 2023). Esse direito torna-se importante de uma maneira substantiva, pois reconhece que o indivíduo está sujeito a decisões automatizadas e processual, porque reconhece que pode haver provas autônomas para reivindicações legais (Dourado & Aith, 2022).

Mas a implementação do Artigo 20 apresenta três problemas específicos, identificados por algumas literaturas especializadas. O primeiro problema diz respeito ao que constitui uma decisão exclusivamente automatizada. Na realidade, sistemas muito complexos raramente são automatizados em completo isolamento da intervenção humana (Laux, 2023). Um estudo recente sobre o Ato de IA Europeu mostra que a tendência operacional em grandes empresas sempre foi de "carimbo de borracha": um humano está formalmente validando decisões algorítmicas em um fluxo contínuo de dados sem realmente processar o conteúdo, o que reduz a proteção legal a uma formalidade vazia (Laux, 2023). Portanto, decisões verdadeiramente automatizadas não parecem estar totalmente cobertas pelo Artigo 20 (Laux, 2023).

A segundo problema refere-se à tensão entre a transparência algorítmica e os segredos comerciais (Dourado & Aith, 2022). A LGPD exige que o controlador forneça informações claras e adequadas sobre os critérios de tomada de decisão (Carvalho et al., 2019); a divulgação da lógica subjacente, muitas vezes protegida por direitos de propriedade intelectual ou segredos industriais, entra em conflito com a proteção legal paralela (Dourado & Aith, 2022).

A doutrina é mais material do que transparente: a necessidade de transparência não é a divulgação do código-fonte, mas uma explicação funcional e contextual que seja suficiente para que o titular compreenda os critérios relevantes e os conteste (Dourado & Aith, 2022; Laux, 2023). Essa abordagem, em linha com o princípio da transparência contextual do AI Act, é intermediária e requer a capacidade de auditoria técnica da ANPD (Camargo, 2024; Sarlet & Rodríguez, 2022).

O terceiro problema é a falta de um mecanismo coercitivo (Laux, 2023). A LGPD permite sanções administrativas genéricas (Santos, 2023), mas não prevê um mecanismo específico para garantir o direito de revisão em um tempo razoável (Alves et al., 2023). Na prática, o titular que se sente prejudicado pela decisão algorítmica deve recorrer ao processo administrativo prévio à ANPD (com um tempo de conclusão superior a dois anos devido ao volume de demandas) ou à proteção civil genérica, com todas as dificuldades probatórias já consideradas (Mulholland, 2025).

Assim, o Artigo 20 até agora serve de orientação interpretativa do que um mecanismo de proteção autônomo (Alves et al., 2023). A solução, enraizada no contexto do Ato Europeu de IA, não é divulgar o código, mas fornecer ao titular informações para entender por que é o algoritmo e desafiá-los. Nesse cenário, a ANPD tem a competência legal para conduzir ou decidir auditorias, que é um mecanismo de monitoramento e garantia de conformidade (Boff; Leal, 2021).

A título de exemplo podemos citar o caso emblemático que ilustra a atuação da Autoridade Nacional de Proteção de Dados ANPD envolvendo a empresa Meta em 2024, no qual a agência suspendeu de forma preventiva, a política de privacidade da Meta. Tal política envolvia o uso de dados de cidadãos brasileiros para treinar sistemas de IA, e o argumento central era o risco concreto e elevado de prejuízos causados a essas pessoas pelo uso de seus dados (Camargo, 2024).

Após uma análise do caso pela ANPD e a apresentação, pela empresa Meta, de um plano de conformidade para se adequar às leis vigentes, o tratamento de dados foi liberado novamente. Contudo, essa liberação foi condicionada à adoção de medidas protetivas específicas. Entre elas, destacaram-se os dados de menores de idade, criação de um canal facilitador para que as pessoas contestem o uso de suas informações e umas garantias de maior transparência quanto aos motivos desse tratamento de dados (Camargo, 2024).

Essa decisão da ANPD deu um passo na adequação do quadro normativo existente (Lei Geral de Proteção de Dados), especialmente quanto à responsabilização diante da evolução das tecnologias de IA e à proteção da privacidade, destacando a importância de as empresas agirem com transparência e respeitarem os direitos dos usuários ao utilizar seus dados para treinar modelos de IA, sob o risco de serem responsabilizadas pela falta de implementação de medidas de segurança adequadas quando o risco for razoavelmente previsível (Souza e al., 2024).

O Artigo 6º, IX da LGPD tem um requisito explícito de não discriminação: os dados não devem ser processados para fins discriminatórios, ilícitos ou abusivos (Santos, 2023). Seu significado operacional, no entanto, reside em grande parte na interpretação (Gonçalves; Barbaresco, 2024; Santos, 2023), e a LGPD não toma nenhuma decisão específica para ser punida por discriminação algorítmica e fornece o regime geral de sanções (Gonçalves; Barbaresco, 2024). Essa lacuna, que é parcialmente preenchida pelo Artigo 42 da LGPD (responsabilidade civil), não preenche completamente o Artigo 6º, IX (Gonçalves & Barbaresco, 2024). Como resultado, a literatura sugere duas linhas interpretativas que expandem seu escopo. A primeira é vê-lo como um elemento vinculante no controle de proporcionalidade em decisões judiciais (Gonçalves; Barbaresco, 2024). A segunda é mais ambiciosa, visando ser uma regra obrigatória para o controle de constitucionalidade sub-legal (Gonçalves; Barbaresco, 2024). Sua fragilidade atualmente não está produzindo efeitos diretamente úteis (Gonçalves; Barbaresco, 2024; Santos, 2023).

O Artigo 42 da LGPD estabelece que o controlador ou operador que, pelo tratamento de dados pessoais, cause dano patrimonial, moral, individual ou coletivo a terceiros em violação de qualquer lei é responsável por repará-lo (Alves et al., 2023). Seu parágrafo 2 também permite a inversão do

ônus da prova (Alves et al., 2023) sobre o titular. Segundo nós, esta disposição é a peça material mais necessária do atual sistema jurídico brasileiro para a proteção das vítimas de discriminação algorítmica, pois é o instrumento material mais útil porque (i) retira, em certa medida, o ônus da prova da vítima (Estêves; João, 2025) (ii) estabelece o controlador como responsável por natureza, em linha com o quadro europeu (Alves et al., 2023) (iii) atua como uma âncora normativa para o princípio da transparência (Alves et al., 2023) na prática (Alves et al., 2023).

Portanto, a fragilidade do Artigo 42 está em seus limites (Estêves; João, 2025; Mulholland, 2025) sendo sua aplicação condicionada ao caso de uma violação da LGPD e não é aplicável a casos de discriminação com base em outros fundamentos normativos; não é uma solução automática no que diz respeito à inversão do ônus da prova, que deve ser determinada por uma avaliação judicial (Alves et al., 2023); e sua articulação com regimes simultaneamente aplicáveis, como o CDC, precisará de construção jurisprudencial (Alves et al., 2023; Silva, 2026).

Além disso, a LGPD tem potencial regulatório para proteger contra a discriminação algorítmica, mas é um instrumento que depende muito da construção jurisprudencial e do desenvolvimento da capacidade institucional da ANPD (Camargo, 2024; Sarlet; Rodríguez, 2022; Souza et al., 2024). O caso Meta é um sinal da tendência positiva (Camargo, 2024) e a necessidade de preencher suas lacunas legislativas (sanções específicas para discriminação algorítmica e prazos para responder a pedidos de revisão) está na agenda (Barbosa, 2025; Mulholland, 2025).

4.3. Código de Defesa do Consumidor: produto defeituoso e inversão do ônus.

O Código de Defesa do Consumidor (Lei nº 8.078/1990) oferece o terceiro eixo normativo que pode ser mobilizado para a proteção de vítimas de discriminação algorítmica através do uso do conceito de “produto ou serviço defeituoso”, a responsabilidade objetiva do fornecedor (art. 14) e a inversão judicial do ônus da prova (art. 6º, VIII) (Júnior, 2025; Paulo, 2022). Sua relevância é particularmente destacada em relações onde um sistema de IA é oferecido como produto ou serviço aos consumidores finais, nos mercados financeiro, de varejo e de recrutamento (Maliha, 2021; Silva, 2026).

Conforme sistematizado por Menezes Júnior (2025), a aplicação do regime consumerista à Inteligência Artificial depende da construção de tese alinhada ao CDC: sistemas de IA que produzem resultados estatisticamente enviesados são juridicamente equiparáveis a produtos defeituosos por falha de concepção (Júnior, 2025; Maliha, 2021). A dependência funcional do sistema em dados enviesados ou arquitetura inadequada é, na verdade, um defeito (Júnior, 2025) e o sistema deve ser totalmente responsável conforme especificado nos artigos 12 e 14 do CDC (Júnior, 2025). É um conceito que rompe com a lógica tradicional de análise de defeitos (Júnior, 2025) ao não exigir que se possa

mostrar um certo problema físico ou técnico, mas que o sistema não oferece a segurança legitimamente esperada, pois o sistema não é seguro. As evidências comportamentais e estatísticas (taxa de erro, desproporcionalidade no tratamento de grupos protegidos, comparação com sistemas equivalentes em uso) a esse respeito são de interesse (Maliha, 2021).

O artigo 7º, parágrafo único, do Código de Defesa do Consumidor (CDC) prevê a responsabilidade solidária entre todos os elos da cadeia de produção e distribuição (Silva, 2026), uma ferramenta particularmente útil para a discriminação algorítmica, pois o dano muitas vezes resulta da interação entre múltiplos atores: desenvolvedor, fornecedor de dados, operador do sistema, usuário final (Silva, 2026). Em sistemas complexos, a determinação precisa do ponto de falha técnica tende a ser materialmente inviável (Zech, 2021); a solidariedade permite que a vítima processe qualquer um dos atores (Silva, 2026).

O artigo 28 do Código de Defesa do Consumidor (CDC) estabelece que, quando há abuso de direito ou confusão patrimonial, a personalidade jurídica não deve ser considerada (Rezende; Oliveira, 2019). Isso se torna relevante quando empresas de Inteligência Artificial formam estruturas corporativas complexas para diluir sua responsabilidade. Além disso, a inversão do ônus da prova, prevista no artigo 6º, VIII, do CDC, uma ferramenta para reconciliar a assimetria técnica entre as partes, pois dá ao fornecedor a oportunidade de provar que o processamento não é deficiente (Silva, 2026).

5. MARCO LEGAL EM CONSTRUÇÃO: OS POJETOS DE LEI EM DISPUTA

5.1. PL 2.338/2023: Avanços e limites

O Projeto de Lei nº 2.338/2023, atualmente em tramitação no Senado Federal, é uma das propostas legislativas relevantes do Brasil para a regulamentação da Inteligência Artificial (Cunha; Nascimento, 2025). O projeto é baseado em risco, identificando sistemas de IA como potenciais causadores de danos e estabelecendo as responsabilidades para os desenvolvedores e operadores com base no nível de dano potencial (Viana; Toscano, 2025). Inspirado na Lei de IA da União Europeia, os níveis de risco são divididos em riscos mínimos, limitados, altos e inaceitáveis, com maior responsabilidade para desenvolvedores e operadores à medida que o nível de risco aumenta (Viana; Toscano, 2025). No caso de sistemas de IA de alto risco, como recrutamento e seleção, concessão de crédito, habitação e segurança pública, o Projeto de Lei 2.338/2023 exige que o requerente tenha registrado e relatado o impacto dos sistemas de IA ao governo, realize auditorias independentes regulares e que o ônus da prova seja retirado do sistema jurídico no caso de responsabilidade civil, conforme o Artigo

37 do Projeto da Lei. O progresso feito pelo PL 2.338/2023 é reconhecido pela doutrina majoritária, que destaca sua validade no problema probatório e a necessidade de fornecer um regime de responsabilidade específico para sistemas de IA de alto risco (Esteve; João, 2025).

Em termos de limites, observa-se que o projeto não discute a responsabilidade por danos difusos de natureza existencial (ou seja, danos que afetam a identidade e o projeto de vida da vítima, transcendendo o âmbito patrimonial e não se encaixando nas categorias tradicionais de dano moral individual (Mulholland, 2025). Além disso, a definição de alto risco é criticada por ser muito rígida em alguns contextos e não em outros, podendo assim ser classificada como de baixo risco e, portanto, sujeita a lacunas regulatórias (Frazão, 2024). A eficácia das obrigações no projeto depende de uma ANPD com recursos técnicos e humanos suficientes para gerenciar esses sistemas complexos, o que é uma questão institucional que a doutrina tem alertado (Sarlet; Rodríguez, 2022).

5.2 PROJETO DE LEI 4/2025 (Reforma do Código Civil): contribuições e limitações

O Projeto de Lei 4/2025, que articula a reforma do Código Civil e incorpora as recomendações da Comissão presidida pelo jurista Flávio Tartuce, propõe atualizar o estatuto para enfrentar desafios atuais como a responsabilidade civil no ambiente digital e a Inteligência Artificial (Junior, 2025). Ele busca incluir a responsabilidade objetiva por danos de sistemas autônomos e preencher lacunas no CDC para relações não consumidoras.

Quanto às limitações do PL 4/2025, vários pesquisadores já apontaram que o projeto não resolve o problema probatório. Mesmo que a culpa seja classificada como uma violação dos deveres de governança, é necessário mostrar à vítima qual dever específico foi violado e como a violação está relacionada ao dano causado e por quê uma tarefa que, na prática, replica o desafio probatório apresentado no projeto. Além disso, o projeto é mais adequado para relações contratuais bilaterais do que para danos causados por plataformas digitais a não usuários ou sistemas de justiça preditiva aplicados a réus.

6. RISCOS DISCRIMINATÓRIOS POR SETOR: ANÁLISE VERTICAL

6.1. Direito do Trabalho

A discriminação algorítmica está presente no Direito do Trabalho em três áreas: recrutamento e seleção, monitoramento de desempenho e demissão. No recrutamento, como na Amazon, onde o sistema automatizado de triagem penalizou mulheres no trabalho, essa ferramenta pode ser usada para filtrar candidatos de acordo com uma determinada categoria de candidatos protegidos (como gênero ou raça). No entanto, a Consolidação das Leis do Trabalho (CLT) não especifica se os algoritmos são

tendenciosos em decisões automatizadas de emprego, e a Lei nº 9.029/1995 não proíbe especificamente a discriminação nas relações de trabalho, mas também não a proíbe.

No monitoramento de desempenho, sistemas de avaliação algorítmica como os utilizados por plataformas de trabalho digital, como aplicativos de entrega e transporte, podem criar dinâmicas de controle e avaliação tendenciosas em relação a certos perfis. O desligamento algorítmico ocorre quando a rescisão do contrato de trabalho é recomendada ou realizada por um sistema automatizado sem qualquer intervenção humana. Nesse contexto, a ausência de regulamentação específica para o desligamento algorítmico no Brasil revela-se uma lacuna significativa, especialmente à luz do princípio constitucional de proteção ao emprego (art. 7º, I, da CF/88) e diante do crescente plataformização das relações de trabalho (Cunha, 2024).

6.2. Direito do Consumidor

No Direito do Consumidor, os sistemas de pontuação de crédito e as práticas de precificação dinâmica são grandes fontes de discriminação algorítmica (Wachter; Mittelstadt; Russell, 2021). A pontuação de crédito tem sido criticada por reproduzir padrões passados de exclusão financeira. No Brasil, os algoritmos de crédito tendem a dar pontuações mais baixas aos residentes de regiões periféricas e pessoas com nomes associados a certos grupos étnicos, independentemente de sua capacidade de pagamento (Vilarino; Vicente, 2020).

A Lei do Cadastro Positivo (Lei Complementar nº 166/2019) não fornece nenhuma maneira de auditar objetivamente o viés discriminatório na pontuação de crédito, e a Lei Geral de Proteção de Dados é aplicada em paralelo (Freitas; Mafini, 2020). A precificação dinâmica, na qual os consumidores são cobrados preços diferentes pelo mesmo produto ou serviço com base em seu perfil algorítmico, faz parte do debate sobre precificação personalizada. Nesse modelo, algoritmos de inteligência artificial estimam a disposição de cada consumidor para pagar, a fim de extrair o máximo de dinheiro, o que aumenta a probabilidade de comportamento predatório e pode prejudicar pessoas vulneráveis com base em dados comportamentais (Pantaja; Leão, 2022).

Quando um algoritmo descobre, a partir de dados comportamentais, que um consumidor pertence a um grupo com menor poder de negociação e lhes oferece preços mais altos com condições menos favoráveis, isso constitui discriminação de preços com base em categorias protegidas, mesmo que nenhuma dessas categorias tenha sido explicitamente incluída no modelo (Frazão; Goetterbauer, 2022). Tal hipótese está dentro da proibição de práticas abusivas prevista no Artigo 39 do Código de

Defesa do Consumidor e pode também levar à responsabilidade objetiva do fornecedor, conforme o Artigo 14 do mesmo estatuto (Silva, 2026).

6.3. Direito Administrativo

A discriminação algorítmica é particularmente problemática no Direito Administrativo devido ao poder coercitivo e à extensão em que as decisões estatais são tomadas (Araújo; Zullo; Torres, 2020). As áreas mais importantes a esse respeito são o Judiciário, a segurança pública e a distribuição de benefícios sociais (Borgesius, 2025). No Judiciário, o ASSIS (Assistente de Suporte à Integração e Inteligência Sistemática) é um sistema de apoio à decisão desenvolvido pelo Tribunal de Justiça do Estado do Rio de Janeiro (TJ-RJ) que pode minar o devido processo legal e reproduzir o viés estrutural no sistema de justiça criminal que historicamente tem se concentrado em pessoas negras e pobres (Santiago, 2025).

O devido processo legal (art. 5º, LIV da CF/88) e a exigência de que as decisões judiciais sejam fundamentadas (art. 93, IX da CF/88) são limitações ao uso de IA em áreas urbanas sem regulamentação clara (Siqueira; Moraes; Santos, 2022). O reconhecimento facial em câmeras de vigilância policial é um tema de sério debate na segurança pública, pois tende a reforçar a criminalização seletiva e a vigilância de comunidades periféricas (Lima; Barreto, 2024). Embora o AI Act da União Europeia proíba explicitamente o uso de reconhecimento facial para vigilância geral em espaços públicos, o Brasil não possui regulamentação sobre o assunto. Em municípios como São Paulo e Salvador, sistemas foram implementados fora do marco legal, o que aumenta a identificação falsa e reforça o racismo estrutural, conforme Buolamwini e Gebru (2018) em relação aos algoritmos. O uso de algoritmos pelo INSS para concessão e revisão de benefícios gerou controvérsia porque as negativas automáticas sem justificativa não são apenas injustas, mas também vão contra a base de dados de baixa alfabetização digital e capacidade de contestação.

CONCLUSÃO

O desenvolvimento da Inteligência Artificial em áreas sensíveis impôs à sociedade brasileira desafios inéditos no campo da responsabilidade civil, especialmente quando decisões automatizadas produzem resultados discriminatórios. Conforme demonstrado ao longo deste trabalho, o ordenamento jurídico pátrio ainda carece de um regime normativo claro para categorizar e enquadrar esses sistemas, razão pela qual doutrina e jurisprudência têm recorrido a princípios constitucionais igualdade, dignidade, não discriminação e privacidade – bem como ao Código Civil, ao CDC e à LGPD, na tentativa de suprir essa lacuna.

Tal abordagem interpretativa, contudo, revela-se insuficiente diante da complexidade técnica dos algoritmos, da opacidade inerente aos modelos de aprendizado profundo e da crescente autonomia

decisória das máquinas. O principal óbice à reparação do dano reside na dificuldade probatória enfrentada pela vítima para comprovar o nexo causal entre o resultado discriminatório e a conduta de um agente específico. Essa barreira inviabiliza a aplicação pura da teoria subjetiva (art. 186 do CC) e impõe a adoção de um sistema misto: responsabilidade objetiva fundada no risco (art. 927, parágrafo único, do CC) conjugada à inversão dinâmica do ônus da prova (art. 373, § 1º, do CPC).

Nesse cenário, a efetiva proteção das vítimas de discriminação algorítmica dependerá da implementação de três medidas concretas. Primeiro, é indispensável o fortalecimento técnico e orçamentário da ANPD, com a atribuição expressa de competência para instituir auditorias algorítmicas obrigatórias em sistemas de alto risco, considerando a atual carência de pessoal especializado na agência. Segundo, impõe-se a inclusão, no PL 2.338/2023, de previsão de seguro civil obrigatório para desenvolvedores e operadores de IA garantindo que a indenização à vítima não fique condicionada à solvência do agente causador do dano. Terceiro, a edição de enunciado jurisprudencial ou súmula vinculante que estabeleça, de forma inequívoca, a inversão do ônus da prova como regra geral nas demandas envolvendo decisões automatizadas, conferindo previsibilidade e segurança jurídica ao julgador e à vítima.

Por fim, reconhece-se que o tema dos danos difusos e existenciais em especial o racismo algorítmico estrutural não foi exaurido neste trabalho, permanecendo como agenda relevante para pesquisas futuras que aprofundem a interface entre o direito antidiscriminatório e a governança algorítmica no Brasil.

REFERÊNCIAS

ALMEIDA, Silvio Luiz de. Racismo estrutural. São Paulo: Pólen, 2019.

ALVES, Cândida Helena Lopes; HALABE, Dannilo Jorge Escorcio; HALABE, Maria. Responsabilidade civil na Lei Geral de Proteção de Dados. São Paulo: Revista dos Tribunais, 2023.

ARAÚJO, Valter Shuenquener de; ZULLO, Bruno Almeida; TORRES, Maurílio. Big data, algoritmos e inteligência artificial na administração pública: reflexões para a sua utilização em um ambiente democrático. Revista de Direito Administrativo, Rio de Janeiro, v. 281, n. 1, p. 95-122, jan./abr. 2020.

BRASIL. Senado Federal. Projeto de Lei nº 2.338, de 2023. Dispõe sobre o uso da inteligência artificial. Brasília, 2023. Disponível em: <https://www.senado.leg.br/>. Acesso em: 08 maio.2026.

BRASIL. Constituição Federal de 1988. Promulgada em 5 de outubro de 1988. Disponível em < http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm> Acesso em: 10 maio.2026.

BRASIL. Lei nº 10.406 (Código Civil), de 10 de janeiro de 2002. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/2002/L10406compilada.htm. Acesso em: 10 maio 2026.

BRASIL. Lei nº 8.078/1990 (Código de Defesa do Consumidor), de 11 de setembro de 1990. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/18078.htm. Acesso em: 10 mai.2026.

- BRASIL. Lei nº 13.105 (Código de Processo Civil), de 16 de março de 2015. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/13105.htm#art373. Acesso em: 10 maio 2026.
- BRASIL. Lei nº 13.709, (Lei Geral de Proteção de Dados Pessoais), de 14 de agosto de 2018. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato20152018/2018/lei/13709.htm. Acesso em: 10 de maio. 2026.
- BRASIL. Lei nº 9.029, (Proíbe a exigência de atestados de gravidez e esterilização, e outras práticas discriminatórias) de 13 de abril de 1995. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/19029.htm. Acesso em: 10.mai.2026.
- BARBOSA, Mafalda Miranda. IA, riscos e responsabilidade – uma reflexão em torno do Regulamento IA e do Projeto Lei brasileiro nº 2338, de 2023. *Revista de Direito Civil, Coimbra*, v. 14, n. 2, p. 45-78, 2025.
- BUOLAMWINI, Joy; GEBRU, Timnit. Gender shades: Intersectional accuracy disparities in commercial gender classification. In: *Conference on fairness, accountability and transparency*. PMLR, 2018. p. 77-91.
- BUITEN, Miriam Caroline. Product liability for defective AI. *European Journal of Risk Regulation*, Cambridge, v. 15, n. 2, p. 245-268, 2024.
- BORGES, Frederik Zuiderveen. Discrimination, artificial intelligence, and algorithmic decision-making. *ArXiv.org*, 15 out. 2025.
- CORREA, Bianca Kremer Nogueira. DIREITO E TECNOLOGIA EM PERSPECTIVA AMEFRICANA Autonomia, algoritmos e vieses raciais. Disponível em: <<https://www.maxwell.vrac.puc-rio.br/58993/58993.PDF>>. Acesso em: 08 jun. 2026.
- CAMARGO, Solano de. ANPD autoriza Meta a retomar uso de dados pessoais para IA: inovação ou risco à privacidade? *Jornal da Advocacia – OAB-SP*, São Paulo, 04 set. 2024. Disponível em: <https://www.oabsp.org.br/jornaldaadvocacia/24-09-02-1548-anpd-autoriza-meta-a-retomar-uso-de-dados-pessoais-para-ia-inovacao-ou-risco-a-privacidade>. Acesso em: 25 março 2026.
- CARDOSO, Oscar Valente. Inteligência Artificial e Políticas Públicas: Desafios Regulatórios e Perspectivas Futuras. *Revista Brasileira de Direito Público*, Belo Horizonte, v. 23, n. 71, p. 89-118, 2025.
- CARVALHO, Luiz Paulo; OLIVEIRA, Jonice; CAPPELLI, Cláudia et al. Desafios de Transparência pela Lei Geral de Proteção de Dados Pessoais. *iSys – Revista Brasileira de Sistemas de Informação*, Rio de Janeiro, v. 12, n. 1, p. 73-97, 2019.
- CUNHA, Fábio Alessandro Neves da; NASCIMENTO, Márcio de Jesus Lima do. Aspectos jurídicos da inteligência artificial no Brasil: comparativo entre o Projeto de Lei 2.338/23 e o Regulamento da União Europeia. *Revista Jurídica da Escola Superior do Ministério Público*, São Paulo, v. 18, n. 35, p. 33-58, 2025.
- DONEDA, Danilo; MENDES, Laura Schertel; SOUZA, Carlos Affonso Pereira de et al. Considerações iniciais sobre inteligência artificial, ética e autonomia pessoal. *Revista de Informação Legislativa*, Brasília, v. 55, n. 218, p. 13-30, abr./jun. 2018.
- DOURADO, Daniel de Araujo; AITH, Fernando. A regulação da inteligência artificial na saúde no Brasil começa com a Lei Geral de Proteção de Dados Pessoais. *Revista de Direito Sanitário*, São Paulo, v. 23, n. 1, p. 89-115, 2022.
- ESTEVES, M.; SÃO JOÃO, Natália. Responsabilidade civil por ato de inteligência artificial, à luz do projeto de lei 2.338, de 2023. *Revista de Direito e Novas Tecnologias*, São Paulo, v. 7, n. 2, p. 67-94, 2025.
- FLORIDI, Luciano. *The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities*. 2023
- FERREIRA, Thiago Amparo; PRADO, Viviane Muller. Racismo creditício no Brasil e nos EUA: risco discriminatório no acesso a crédito. *Revista Direito GV*, São Paulo, v. 20, e 2408, p. 1-32, 2024.

- FRAZÃO, Ana. Regulation of artificial intelligence in Brazil: examination of Draft Bill no. 2338/2023. *Journal of Law, Technology & the Internet*, Cambridge, v. 5, n. 2, p. 87-114, 2024.
- FRAZÃO, Ana; GOETTENAUER, Carlos. Algoritmos de precificação e Direito Concorrencial. *Revista do IBRAC – Instituto Brasileiro de Estudos de Concorrência, Consumo e Comércio Internacional*, São Paulo, v. 38, n. 1, p. 113-148, 2022.
- FREITAS, Cinthia Obladen de Almendra; MAFFINI, Maylin. A proteção dos dados pessoais no crédito bancário e a Lei Geral de Proteção de Dados frente ao cadastro positivo. *Revista de Direito do Consumidor*, São Paulo, v. 124, p. 215-247, jul./ago. 2020.
- FUSTER, Andreas; GOLDSMITH-PINKHAM, Paul; RAMADORAI, Tarun et al. Predictably unequal? The effects of machine learning on credit markets. *The Journal of Finance*, New York, v. 77, n. 1, p. 5-47, Feb. 2022.
- GONÇALVES, Rafaela Vilela; BARBARESCO, Rogério Ananias. Discriminação em algoritmos de inteligência artificial: estudo da LGPD como mecanismo de controle dos vieses discriminatórios. *Revista de Direito e Novas Tecnologias*, São Paulo, v. 6, n. 2, p. 45-78, 2024.
- JOHNSON, Gabbrielle M. Algorithmic bias: on the implicit biases of social technology. *Synthese*, Amsterdam, v. 198, n. 10, p. 9941-9961, Oct. 2021.
- JÚNIOR, Charles Cunha Menezes. Responsabilidade civil e a inteligência artificial: desafios e perspectivas. *Revista de Direito Civil Contemporâneo*, São Paulo, v. 12, n. 38, p. 89-124, out./dez. 2025.
- LAUX, Johann. Institutionalised distrust and human oversight of artificial intelligence: towards a democratic design of AI governance under the European Union AI Act. *Regulation & Governance*, Hoboken, v. 17, n. 4, p. 1023-1042, Oct. 2023.
- MALIIHA, George; SHERIDAN, Genevieve; WEISLEDER, Eric A. et al. Artificial intelligence and liability in medicine: balancing safety and innovation. *Milbank Quarterly*, New York, v. 99, n. 3, p. 629-647, Sept. 2021.
- MAALEJ, Nestor; FATTORINI, Loredana; PERRAULT, Raymond et al. Artificial intelligence index report 2024. Stanford: Stanford Institute for Human-Centered Artificial Intelligence, 2024. Disponível em: <https://aiindex.stanford.edu/report/>. Acesso em: 15 mai. 2025.
- MATEUS, Cátia. Como a Amazon foi atraída pelo algoritmo sexista. *Público*, Lisboa, 11 out. 2018. Disponível em: <https://www.publico.pt>. Acesso em: 12 mar. 2025.
- MULHOLLAND, Caitlin. Responsabilidade civil pelos danos ocasionados pelo uso de inteligência artificial: uma crítica à redação atual do Projeto de Lei nº 2.338/2023. *Revista Brasileira de Direito Civil Rio de Janeiro*, v. 18, n. 105, p. 67-96, jul./set. 2025.
- OBERMEYER Ziad; POWERS Brian; VOGELI Christine, MULLAINATHAN Sendhil. Analisando o viés racial em um algoritmo usado para gerenciar a saúde de populações. *Science* 366 , 447-453 (2019). DOI: 10.1126/science.aax2342. Disponível em: <https://www.science.org/doi/10.1126/science.aax2342> Acesso em: 13 de maio.2026.
- PANTOJA, Adilson Carvalho; LEÃO, Semírames de Cássia Lopes. Discriminação algorítmica no mercado de trabalho: o viés preditivo na interpretação de dados. *Revista do Tribunal Superior do Trabalho*, Brasília, v. 88, n. 4, p. 113-138, out./dez. 2022.
- PAULO, Arsénio. Por uma terceira via reforçada a partir de um poder legislativo forte e ético. *Revista da Faculdade de Direito da UFMG*, Belo Horizonte, n. 78, p. 145-180, jul./dez. 2022.
- RABAY, Gustavo; SILVA, Claudyvan; SILVA, Joice. Discriminação social para além do ciberespaço: proteção de direitos na era digital – protection of legal subjects in the framework of algorithmic governance and the global expansion of artificial intelligence applications. *Revista Direito e Justiça*, Porto Alegre, v. 47, n. 2, p. 78-105, 2025.
- RODRIGUES, Leandro Nascimento da Silva; LEAL, Pastora do Socorro Teixeira. A eficácia dos direitos fundamentais nas relações privadas à luz da jurisprudência do STF: análise crítica do RE

- 201.819-8 e ADI 4815. *Revista de Direitos Fundamentais & Democracia*, Curitiba, v. 23, n. 2, p. 89-115, 2018.
- RODRIGUES, Rowena. Legal and human rights issues of AI: gaps, challenges and vulnerabilities. *Computer Law & Security Review*, Amsterdam, v. 38, 105408, p. 1-15, Sept. 2020.
- SANTOS, Vanessa Cristina Bissoli dos. The General Data Protection Law – Law no. 13709/2018, and compliance with the Principle of Non-Discrimination. São Paulo: Editora Unesp Digital, 2023.
- SARLET, Gabrielle Bezerra Sales; RODRÍGUEZ, Daniel. Autoridade Nacional de Proteção de Dados: análise da sua estrutura, atribuições e dos desafios de atuação. *Revista de Direito Constitucional e Internacional*, São Paulo, v. 30, n. 123, p. 145-178, out./dez. 2022.
- SELENA, Silva; KENNEY, Martín. Algorithms, platforms, and ethnic bias: an integrative essay. *Sociology of Race and Ethnicity*, Washington, v. 5, n. 3, p. 399-412, July 2019.
- SILVA, Vitória Soares da. A responsabilidade civil das plataformas de inteligência artificial generativa sob a ótica do Código de Defesa do Consumidor e a proteção de dados. *Revista de Direito Civil Contemporâneo*, São Paulo, v. 13, n. 42, p. 145-180, jan./mar. 2026.
- SIQUEIRA, Dirceu Pereira; MORAIS, Fausto Santos de; SANTOS, Marcel Ferreira dos. Inteligência artificial e jurisdição: dever analítico de fundamentação e os limites da substituição dos humanos por algoritmos no campo da tomada de decisão judicial. *Revista Brasileira de Direito Processual Penal*, Porto Alegre, v. 8, n. 3, p. 1455-1492, set./dez. 2022.
- SOUZA, Alcian Pereira de; SIMAS, Danielle Costa de Souza; JUSTINIANO, Jeibson dos Santos et al. O papel da ANPD na defesa da proteção de dados: análise da suspensão do tratamento de dados pessoais para o treinamento da inteligência artificial pela Meta. *Revista Síntese – Direito Civil e Processual Civil*, São Paulo, v. 32, n. 102, p. 89-115, mar./abr. 2024.
- SULLIVAN, Hannah; SCHWEIKART, Scott J. Are current tort liability doctrines adequate for addressing injury caused by AI? *American University Law Review*, Washington, v. 68, n. 1, p. 191-244, 2019.
- SCHERTEL, Mendes, L., & Mattiuzzo, M. (2019). Discriminação algorítmica: conceito, fundamento legal e tipologia. *Direito Público*, 16(90). Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/3766>. Acesso em: 14 mai. 2026.
- VIANA, Henrique Silviano Almeida; TOSCANO, Stefano. Breve análise comparativa entre o Projeto de Lei n. 2.338/2023, do Brasil e o Regulamento 2024/1689: semelhanças e divergências à luz da classificação dos riscos das atividades a serem desempenhadas por IA. *Revista de Direito Internacional e Globalização Econômica*, São Paulo, v. 14, n. 27, p. 67-104, 2026.
- VIEIRA, Ana Elisa Silva Fernandes; SIQUEIRA, Dirceu Pereira. Reflexões sobre o controle algorítmico diante o controle do direito formal e os riscos aos direitos da personalidade. *Revista de Direito e Novas Tecnologias*, São Paulo, v. 5, n. 3, p. 67-94, set./dez. 2023.
- VILARINO, Ramon Casas; VICENTE, Renato. An experiment on the mechanisms of racial bias in ML-based credit scoring in Brazil. *Proceedings of the 1st ACM Workshop on Algorithmic Fairness in Finance*, New York: ACM, 2020. p. 1-6.
- WACHTER, Sandra; MITTELSTADT, Brent; RUSSELL, Chris. Bias preservation in machine learning: the legality of fairness metrics under EU non-discrimination law. *SSRN Electronic Journal*, Rochester, NY, 31 Dec. 2021. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3792775. Acesso em: 12 abr. 2025.
- WANG, Xukang; WU, Ying Cheng; JI, Xueliang et al. Algorithmic discrimination: examining its types and regulatory measures with emphasis on US legal practices. *AI and Ethics*, Cham, v. 4, n. 3, p. 913-927, Aug. 2024.
- WENDEHORST, Christiane. Strict liability for AI and other emerging technologies. *Journal of European Tort Law*, Berlin, v. 11, n. 2, p. 171-204, Aug. 2020.
- ZECH, Herbert. Liability for AI: public policy considerations. *ERA Forum*, Berlin, v. 22, n. 1, p. 21-29, Apr. 2021.

ZOU, James; SCHIEBINGER, Londa. AI can be sexist and racist – it's time to make it fair. *Nature*, London, v. 559, n. 7714, p. 324-326, July 2018.

ZUIDERVEEN BORGESIU, Frederik. Discrimination, artificial intelligence, and algorithmic decision-making. Strasbourg: Council of Europe, 2025. Disponível em: <https://www.coe.int/en/web/algorithms-and-human-rights>. Acesso em: 17 mar. 2025.