INSTITUTO BRASILEIRO DE ENSINO, DESENVOLVIMENTO E PESQUISA
CURSOS DE PÓS-GRADUAÇÃO *LATO SENSU*

DISCRIMINAÇÃO ALGORÍTMICA EM SEGURANÇA PÚBLICA: ESTUDO DE CASO SOBRE OS SISTEMAS DO MINISTÉRIO DA JUSTIÇA E SEGURANÇA PÚBLICA - MJSP

Márcio Pereira Lima[1]

Tainá Aguiar Junquilho[2]

RESUMO

O presente estudo aborda a discriminação algorítmica no âmbito da segurança pública no Brasil, explorando como o uso de sistemas de Inteligência Artificial pelo setor público pode perpetuar preconceitos e discriminações. O estudo combina pesquisa bibliográfica com uma análise prática, incluindo a busca por informações no site do Ministério da Justiça e Segurança Pública (MJSP), no Portal de Dados Abertos e pedido de acesso à informação realizado em 04 jun. 2025. O uso de IA e *Big Data* na segurança pública, embora prometa eficiência, acarreta um risco potencial de discriminação, causado principalmente por vieses nos dados de treinamento, que podem ser incompletos ou não representativos, reproduzindo desigualdades sociais, prisões arbitrárias e violação de direitos fundamentais. Uma das maiores preocupações identificadas é a falta de transparência sobre o funcionamento dos algoritmos. A pesquisa revela que, apesar da existência de sistemas que utilizam IA no MJSP, há uma carência de transparência dos algoritmos, dificultando a identificação e correção de vieses. O artigo propõe uma abordagem intermediária para mitigar a discriminação algorítmica: em vez da transparência total, que poderia comprometer operações policiais, sugere a divulgação de um conjunto mínimo de informações sobre o funcionamento dos algoritmos. Além disso, ressalta a importância da supervisão humana, de auditorias independentes, do treinamento das equipes de desenvolvedores e garantia da diversidade social nessas equipes. O estudo reforça a importância da implementação de uma governança de IA robusta, para que sirva à justiça e assegure os direitos fundamentais para todos.

PALAVRAS-CHAVE: Discriminação algorítmica; inteligência artificial; segurança pública; Ministério da Justiça e Segurança Pública; governança de IA; transparência algorítmica.

ABSTRACT

This study addresses algorithmic discrimination in the context of public security in Brazil, exploring how the use of Artificial Intelligence (AI) systems by the public sector can perpetuate prejudice and discrimination. The study combines a literature review with a practical analysis, including searching for information on the website of the Ministry of Justice and Public Security (MJSP), the Open Data Portal, and an information request made on June 04, 2025. While the use of AI and Big Data in public security promises efficiency, it carries a potential risk of discrimination, caused primarily by biases in training data. This data can be incomplete or unrepresentative, reproducing social inequalities, leading to arbitrary arrests, and violating fundamental rights. One of the main concerns identified is the lack of algorithmic transparency. The research reveals that despite the existence of AI systems within the Ministry of Justice and Public Security, there is a lack of transparency regarding their algorithms, making it difficult to identify and correct biases. The article proposes an intermediate approach to mitigate algorithmic discrimination: instead of full transparency, which could compromise police operations, it suggests the disclosure of a minimum set of information about how the algorithms work. Furthermore, it highlights

[1] *Servidor público federal, bacharel em Direito e em Análise de Sistemas, mestre em Mecatrônica e especializando em Direito Digital, Dados e Inteligência Artificial, http://lattes.cnpq.br/5822556413524231.*

[2] *Advogada, CEO, Pesquisadora e Professora de Direito, Inovação e Tecnologia no mestrado do IDP, doutora em Direito com ênfase em Inteligência Artificial pela Universidade de Brasília, http://lattes.cnpq.br/5848504606151120.*

the importance of human review, independent audits, training for development teams, and ensuring social diversity within these teams. The study reinforces the importance of implementing robust AI governance to ensure it serves justice and secures fundamental rights for all.

KEYWORDS: *Algorithmic discrimination; artificial intelligence; public security; Ministry of Justice and Public Security; AI governance; algorithmic transparency.*

1 - INTRODUÇÃO

O avanço da Tecnologia da Informação – TI – e da Inteligência Artificial[3] – IA – tem afetado e transformado diversas áreas da sociedade, incluindo o setor público, impactando significativamente a prestação de serviços públicos, promovendo agilidade, eficiência, redução de custos, transparência e participação na formulação de políticas públicas.

Entre os serviços públicos oferecidos pelo Estado à sociedade está a segurança pública, que também busca modernizar e aperfeiçoar a sua atuação.

A segurança pública, conforme disposto no art. 144 da Constituição Federal, é "dever do Estado, direito e responsabilidade de todos, é exercida para a preservação da ordem pública e da incolumidade das pessoas e do patrimônio [...]" (BRASIL, 1988). Sua concretização envolve atuação coordenada de diversos órgãos, como as polícias federal, civis e militares, e da participação ativa de cada cidadão.

Para implementar as ações de segurança pública, os órgãos têm avançado no uso de tecnologias com IA e *Big Data*[4], principalmente para aperfeiçoar a investigação, a prevenção e o combate à criminalidade (BRASIL, 2019; JUNQUILHO e DIAS, 2024, p. 140).

Contudo, a utilização de dessas tecnologias pode gerar problemas quando ocorrem vieses e discriminação, prejudicando indivíduos de determinadas classes sociais, raças, gêneros, entre outras, principalmente as minorias. Isso ocorre principalmente devido ao treinamento de sistemas de IA com bases de dados insuficientemente representativas. No âmbito da segurança pública esse problema é ainda mais grave, devido às consequências para as pessoas

[3] "A Inteligência Artificial pode ser definida como um sistema de ação racional projetado para aprender e se adaptar com base em dados e experiências, com o objetivo de reconhecer padrões, fazer previsões ou tomar decisões" (SAINZ *et. al*, 2024, p. 259).

[4] "*Big Data* é um tipo de conjunto de dados usado em *analytics* que possui imensa quantidade de dados variados e complexos que não podem ser processados por sistemas de gerenciamento de dados tradicionais" (IBM, 2025).

investigadas, como prisões arbitrárias e violação de direitos fundamentais (BRASIL, 2021a, p. 46; NUNES, 2022; JUNQUILHO e DIAS, 2024, p. 127).

Nesse contexto, o Ministério da Justiça e Segurança Pública – MJSP – desempenha um papel central na formulação, coordenação e execução das políticas de segurança pública no Brasil. Embora a segurança pública seja uma responsabilidade compartilhada entre todos os entes da federação, o MJSP atua na vanguarda como órgão central do Sistema Único de Segurança Pública – Susp – sendo responsável por grande parte das ações e diretrizes de segurança pública no país (BRASIL, 2018).

Assim, este artigo tem como objetivo analisar os impactos da discriminação algorítmica em sistemas de segurança pública no âmbito do MJSP, dada a sua relevância estratégica, verificar eventuais medidas implementadas que mitiguem o problema, e contribuir com uma proposta de solução para ampliação da transparência que considere as boas práticas de governança algorítmica e as particularidades inerentes aos serviços de segurança pública.

Ao delimitar o estudo ao MJSP, o artigo ganha relevância prática, pois o ministério é o órgão central na formulação e coordenação das políticas de segurança pública no país, tornando a análise de seus sistemas fundamental para o debate nacional.

Para isso, buscou-se identificar os sistemas de IA e *Big Data* que são utilizados pelo MJSP em atividades de segurança pública, seu estágio de operacionalização, o nível de transparência do seu funcionamento e os eventuais impactos discriminatórios sobre os cidadãos relativos a direitos fundamentais.

A justificativa e a relevância deste trabalho residem em sua abordagem crítica e propositiva sobre um tema da atualidade e de grande impacto social e jurídico: o uso de IA na segurança pública brasileira frente ao seu potencial discriminatório. Há uma premente necessidade de investigar e mitigar os riscos inerentes ao uso de IA nesse setor, buscando equilibrar a busca por eficiência estatal com a proteção dos direitos fundamentais das pessoas.

A metodologia adotada neste trabalho considerou a pesquisa aplicada, incluindo pesquisa bibliográfica com um estudo de caso, dirigido aos sistemas de IA do MJSP. A pesquisa bibliográfica considerou os principais normativos de regência do tema de IA aplicada ao setor público brasileiro, além de referências sobre discriminação algorítmica. Além disso, foi realizada extensa busca no portal do MJSP, no Portal de Dados Abertos e um pedido de acesso à informação junto ao órgão, no intuito de obter informações sobre sistemas de IA do MJSP aplicados à segurança pública.

SGAS Quadra 607 - Módulo 49
Via L2 Sul, Brasília - DF
CEP 70.200-670

(61) 3535-6565

Desse modo, a estrutura deste estudo aborda a discriminação algorítmica da seguinte forma: inicialmente, apresenta-se o conceito, as ocorrências emblemáticas, as causas e as consequências desse fenômeno em sistemas de IA na área de segurança pública no Brasil e no mundo. Na sequência, são apresentadas as principais normas brasileiras que tratam de IA, assim como orientações elaboradas por órgãos públicos para o seu uso de forma responsável e ética. Em seguida, o estudo apresenta os resultados da pesquisa sobre sistemas de informação que utilizam IA e *Big Data* para apoiar as atividades de segurança pública no âmbito do MJSP, avaliando também eventual transparência algorítmica desses sistemas. Posteriormente, são apresentadas formas de mitigação da discriminação algorítmica e promoção da governança algorítmica. Por fim, o estudo é concluído com um resumo dos resultados e uma proposta de solução para mitigar a discriminação em operações de segurança pública que utilizam IA.

2 – DISCRIMINAÇÃO ALGORÍTMICA

Antes de abordar o tema de discriminação algorítmica em si, é fundamental estabelecer as bases constitucionais e legais do combate à discriminação. A partir dessa contextualização será possível aprofundar a análise de como a discriminação pode se manifestar nos sistemas de IA empregados em atividades de segurança pública.

2.1 – DISCRIMINAÇÃO

A discriminação, no seu sentido geral, refere-se a qualquer tipo de distinção baseada em características como gênero, etnia, raça, idade, orientação sexual, nacionalidade, língua, religião, condição social entre outros (ONU, 1948).

No Brasil, a Constituição Federal veda a discriminação como um de seus objetivos fundamentais[5] (BRASIL, 1988). Além do texto constitucional, diversas leis brasileiras proíbem o tratamento discriminatório das pessoas, como: Lei nº 7.716/1989 – Lei Caó –, que define os crimes resultantes de preconceito de raça, cor, etnia, religião ou procedência nacional; Lei nº 12.288/2010 – Estatuto da Igualdade Racial –, que visa garantir à população negra a efetivação da igualdade de oportunidades, a defesa de seus direitos e o combate à discriminação; Lei nº 8.069/1990 – Estatuto da Criança e do Adolescente –, que proíbe qualquer forma de discriminação contra crianças e adolescentes; Lei nº 13.146/2015 – Estatuto da Pessoa com Deficiência –, que assegura direitos e proíbe a discriminação contra pessoas com deficiência; e

[5] "Art. 3º Constituem objetivos fundamentais da República Federativa do Brasil: [...] IV - promover o bem de todos, sem preconceitos de origem, raça, sexo, cor, idade e quaisquer outras formas de discriminação".

a Lei nº 10.741/2003 – Estatuto do Idoso –, que veda a discriminação da pessoa idosa.

No âmbito da pesquisa jurídica, Saiz *et al.* (2024, p. 267), indicam que a discriminação algorítmica se manifesta em pelo menos três eixos temáticos no Direito brasileiro, relacionados com: I) os impactos da IA no sistema judicial; II) as questões de proteção de direitos humanos e direitos fundamentais; e III) os aspectos de transparência, ética e regulamentação da IA no Brasil. Este artigo foca nos eixos II e III, aplicados às políticas de segurança pública no Brasil.

Um estudo sobre discriminação algorítmica em publicações no Brasil até 2024, elaborado por Saiz *et al.* (2024, p. 268), evidencia que 75% delas apresentam exemplos práticos de discriminação, sendo a discriminação racial o tipo mais citado, com 24 ocorrências.

Incidentes envolvendo tecnologias de reconhecimento facial são frequentemente citados nos exemplos de racismo algorítmico (NUNES, 2022; JUNQUILHO e DIAS, 2024; SAIZ *et al.*, 2024), como o caso do Google que categorizou erroneamente imagens de dois jovens negros como gorilas (BBC, 2015).

Nunes (2022) explica que o racismo algorítmico é uma forma de discriminação sistemática e estrutural manifestada e replicada por algoritmos[6] e sistemas de IA, sendo uma consequência direta dos vieses presentes nos dados utilizados para treinar esses sistemas. No caso de tecnologias de reconhecimento facial, os algoritmos são treinados com bancos de dados populados majoritariamente com faces de pessoas brancas, associando-as como humanas. Tal desequilíbrio demográfico nos dados de treinamento leva à sub-representação da população negra e de grupos minoritários, comprometendo o correto funcionamento do algoritmo para reconhecer faces de indivíduos que não pertencem ao grupo majoritário (NUNES, 2022).

Além do racismo algorítmico evidente nas tecnologias de reconhecimento facial, Silva (2022) destaca manifestações mais sutis que permeiam plataformas digitais, como em: sistemas de recomendação de conteúdo, moderação, motores de busca e processamento de imagens.

Essas práticas algorítmicas provocam microagressões que, em última instância, resultam na criminalização presumida, na inferência de inferioridade intelectual, na marginalização cultural, na exclusão e no isolamento de pessoas negras (SILVA, 2022).

A discriminação de gênero emerge como o segundo tipo mais frequente de viés algorítmico documentado. Um caso emblemático é o do *chatbot* Tay, da Microsoft. Este sistema rapidamente internalizou e replicou comportamentos misóginos, transfóbicos e xenófobos,

[6] Algoritmo computacional é "uma série de etapas para completar uma tarefa que é descrita de maneira precisa o bastante para que um computador possa realizá-la" (MENDES e MATTIUZZO, 2019, p. 41).

refletindo os preconceitos presentes nos dados com os quais foi treinado (VINCENT, 2016).

Outro exemplo notório ocorreu na empresa Amazon, em que um algoritmo de recrutamento penalizava a pontuação de candidatas mulheres. Esse viés surgiu pois o modelo foi treinado com base em currículos anteriores, que eram predominantemente de homens, perpetuando assim a disparidade de gênero existente no processo seletivo (DASTIN, 2018).

No que se refere à discriminação geográfica, a prática mais conhecida é a *geopricing*, que aplica preços diferenciados de produtos ou serviços com base na localização geográfica do consumidor. Um caso notório ocorrido no Brasil envolveu a empresa Decolar, que aplicava preços diferenciados e, em alguns casos, inviabilizava reservas a usuários com base em sua geolocalização. Essa conduta levanta questionamentos sobre a potencial associação da localização do indivíduo a condições econômicas menos favoráveis, resultando em práticas discriminatórias (VEJA, 2018).

Esses casos ilustram como a discriminação pode se manifestar em sistemas algorítmicos, com consequências significativas em diversos contextos.

2.2 – A TENDÊNCIA INERENTE À DISCRIMINAÇÃO EM SISTEMAS DE INTELIGÊNCIA ARTIFICIAL

O desenvolvimento tecnológico tem viabilizado o processamento massivo de dados com mais eficiência, permitindo novas correlações entre dados para a produção de previsões de comportamentos futuros (MENDES e MATTIUZZO, 2019).

A despeito de sua relevância, o uso dessas tecnologias traz questões de natureza ética e legal que precisam ser consideradas (JUNQUILHO e DIAS, 2024).

Com a crescente presença da Tecnologia da Informação e da IA no cotidiano e em processos decisórios, essas tecnologias tornam-se potenciais vetores de amplificação de discriminação, casos seus algoritmos estejam enviesados, já contendo padrões discriminatórios e sem tratamento adequado para correções (MENDES e MATTIUZZO, 2019).

Segundo Mendes e Mattiuzzo (2019, p. 42), o objetivo primordial de um algoritmo é fornecer respostas com base nos dados de entrada e a assertividade das respostas está diretamente relacionada com a quantidade e a qualidade dos dados disponibilizados. Nesse contexto, é comum a aplicação de algoritmos em grandes volumes de dados – *Big Data* –, que são coletados com facilidade devido ao aumento do poder computacional das tecnologias atuais.

Ainda nesse contexto, a Inteligência Artificial ganha relevância, com destaque para a

área de aprendizagem de máquina – *machine learning* – em que são fornecidos à máquina tanto os dados de entrada quanto os resultados esperados, e o produto do processamento é um algoritmo, capaz de tornar a relação entre dado e resultado verdadeira (DOMINGOS, 2015).

Contudo, os dados coletados e utilizados em sistemas de IA e de *Big Data* podem ser imprecisos e incompletos, e quando utilizados para processamento em larga escala, aumentam as chances de equívocos que podem gerar discriminação (MENDES e MATTIUZZO, 2019), reproduzindo de forma automatizada as desigualdades e discriminações sociais pré-existentes na estrutura social, prejudicando mais os grupos minoritários e historicamente discriminados (MENDES e MATTIUZZO, 2019; JUNQUILHO e DIAS, 2024; NUNES, 2022).

Quanto aos tipos de discriminação algorítmica, Mendes e Mattiuzzo (2019, p. 51-53) classificam em quatro categorias: I) discriminação por erro estatístico, quando os dados são coletados de forma incorreta ou quando há problemas no código do algoritmo que processa os dados de forma incorreta; II) discriminação por generalização, quando um indivíduo é classificado equivocadamente em determinado grupo; III) discriminação por uso de informações sensíveis, quando a discriminação é gerada a partir de dados como religião, sexualidade e biométricos; e IV) discriminação limitadora do exercício de direitos, quando os resultados gerados pelos algoritmos afetam demasiadamente a algum direito.

Ao se tratar de sistemas de IA aplicados em segurança pública, potencialmente todas essas categorias de discriminação algorítmica podem ocorrer, conforme apresentado na próxima seção.

2.3 – SISTEMAS DE IA EM SEGURANÇA PÚBLICA E POTENCIAL DISCRIMINATÓRIO

Quando se trata de utilização de ferramentas de IA em segurança pública e justiça, os Estados Unidos (EUA) têm se destacado, principalmente nas áreas de policiamento preditivo e avaliação de risco de reincidência, para subsidiar decisões judiciais sobre prisão preventiva, sentenciamento e liberdade condicional (COELHO, 2024).

Entre os exemplos proeminentes de racismo algorítmico frequentemente discutidos na literatura, destaca-se o caso estadunidense do COMPAS – *Correctional Offender Management profiling for Alternative Sanctions* –, cujo algoritmo demonstrou viés racial ao classificar réus negros com maior probabilidade de cometer crimes futuros em comparação com réus brancos (BRASIL, 2021a; SILVA, 2022).

SGAS Quadra 607 - Módulo 49
Via L2 Sul, Brasília - DF
CEP 70.200-670

(61) 3535-6565

Outro uso comum da IA na segurança pública no Brasil e no mundo é em sistemas de reconhecimento facial, para identificar casos de violência urbana, foragidos da polícia e pessoas desaparecidas, como as soluções Smart Sampa[7] e Muralha Digital[8], implantadas respectivamente em São Paulo e Curitiba (JUNQUILHO e DIAS, 2024; SÃO PAULO, 2025).

Como exemplo, a solução Smart Sampa contribuiu para a captura de 2.202 foragidos, 1.885 prisões em 2025, localização de 107 pessoas que estavam desaparecidas e 3.265 prisões em flagrante[9] (SÃO PAULO, 2025).

Contudo, a aplicação da IA em segurança pública nem sempre é eficaz e efetiva, havendo críticas devido a casos de perfilamento racial e de falsos positivos, que resultam em falsas identificações de pessoas (BRASIL, 2021a; JUNQUILHO e DIAS, 2024).

No Brasil, já houve prisões indevidas e constrangimentos decorrentes de falhas em sistemas de reconhecimento facial: (i) uma mulher negra foi presa em 2023 durante o evento Pré-Caju em Sergipe, identificada erroneamente pelo sistema de reconhecimento facial como foragida da Justiça[10]; (ii) um homem negro foi abordado em uma festa junina em Salvador, em 2023, e encarcerado injustamente por 26 dias sob a alegação de ter cometido roubo, o sistema de vigilância havia apurado 95% de semelhança entre ele e a pessoa que deveria ser presa[11]; (iii) um aposentado de 80 anos foi identificado pelo sistema Smart Sampa como um estuprador foragido, em 2024, tendo sido liberado somente após a confirmação do erro[12]; e (iv) uma servidora pública negra, que participava da Conferência Estadual de Igualdade Racial no Rio de Janeiro em 2024, foi erroneamente identificada por reconhecimento facial como foragida da justiça[13], evidenciando a persistência de falsos positivos em contextos de segurança pública.

Nunes (2022) destaca que o viés algorítmico em reconhecimento facial utilizado para a segurança pública acarreta mais prisões e abordagens contra a população negra, fato que leva à reflexão da utilidade das tecnologias de reconhecimento facial para a segurança pública.

Percebe-se, portanto, a gravidade do potencial discriminatório que pode ser gerado por

[7] Mais informações em: https://smartsampa.prefeitura.sp.gov.br/. Acesso em: 26 out. 2025.

[8] Mais informações em: https://www.curitiba.pr.gov.br/noticias/prefeitura-de-curitiba-usa-a-ia-para-aumentar-a-seguranca-melhorar-servicos-e-conversar-com-o-cidadao/76247. Acesso em: 23 ago. 20225.

[9] Dados coletados em 26 out. 2025. Mais informações em: https://smartsampa.prefeitura.sp.gov.br/

[10] Mais informações em: https://www.youtube.com/watch?v=pkNPiOBQ9s8. Acesso em 30 nov. 2025.

[11] Mais informações em: https://www.geledes.org.br/com-mais-de-mil-prisoes-na-ba-sistema-de-reconhecimento-facial-e-criticado-por-racismo-algoritmico-inocente-ficou-preso-por-26-dias/ . Acesso em: 30 nov. 2025.

[12] Mais informações em: https://noticias.uol.com.br/cotidiano/ultimas-noticias/2025/04/13/reconhecimento-facial-de-sp-confunde-idoso-com-estuprador-foragido.htm. Acesso em: 30 nov. 2025.

[13] Mais informações em: https://oantagonista.com.br/brasil/reconhecimento-facial-sistema-da-policia-falha-e-confunde-servidora-publica-com-foragida/. Acesso em 30 nov. 2025.

meio de algoritmos ao reproduzirem discriminações sociais pré-existentes, que acabam recaindo sobre grupos historicamente excluídos ou marginalizados (NUNES, 2022; SAIZ *et al.*, 2024, p. 274). Ao invés de combater, os algoritmos podem acabar reforçando resultados discriminatórios (MENDES e MATTIUZZO, 2019).

A principal causa desse fenômeno é o uso de bases de dados incompletas, enviesadas e não representativas para o algoritmo de treinamento dos sistemas de IA, que perpetuam vieses raciais e socioeconômicos, quando não integradas com controles corretivos (COELHO, 2024; JUNQUILHO e DIAS, 2024).

Mendes e Mattiuzzo (2019) e Saiz *et al.* (2024) apontam outra causa para o fenômeno: a obscuridade dos processos decisórios constantes nos algoritmos. A falta de transparência das regras e fluxos de dados que são utilizados nos algoritmos levam à desconfiança e descrédito em sua utilização, dada a dificuldade em afirmar se alguma discriminação ocorreu ou não.

Além disso, os controladores de dados podem predefinir as correlações, transmitindo aos algoritmos os mesmos vieses presentes nos processos tradicionais de tomada de decisões (MENDES e MATTIUZZO, 2019).

Assim, o potencial discriminatório dos algoritmos, a falta de transparência e a ausência de regulação dos algoritmos em sistemas de IA constituem os principais pontos de preocupação dos pesquisadores da área (SAIZ *et al*., 2024, p. 273).

3 – NORMAS BRASILEIRAS E O COMBATE À DISCRIMINAÇÃO ALGORÍTMICA

Além das disposições gerais sobre a vedação de discriminação na legislação brasileira, existem iniciativas específicas para combater e mitigar a discriminação algorítmica, que incluem projetos de lei, decretos e guias de boas práticas para o uso ético e responsável da IA.

3.1 - PROJETO DE LEI Nº 2.338/2023

Em tramitação no Congresso Nacional, o Projeto de Lei nº 2.338/2023 visa dispor sobre "o desenvolvimento, o fomento e o uso ético e responsável da inteligência artificial com base na centralidade da pessoa humana". Atualmente[14] aguarda a realização de audiência pública na Comissão Especial sobre Inteligência Artificial (BRASIL, 2025a).

Entre seus princípios, destacam-se a não discriminação ilícita ou abusiva, a

[14] O acompanhamento da tramitação do Projeto de Lei nº 2.338/2023 pode ser realizado por meio do *link*: https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2487262.

transparência, a explicabilidade, a proteção dos direitos e garantias fundamentais, a prestação de contas, a responsabilização com a reparação integral de danos, a prevenção, a precaução, a mitigação de riscos e danos, e a governança transparente, participativa e orientada à proteção de direitos fundamentais individuais, sociais, coletivos e econômicos.

O princípio da explicabilidade ganha especial destaque, uma vez que reforça o direito das pessoas que tiveram seus dados processados por algoritmos em acessá-los e compreender como foram utilizados na decisão automatizada.

O PL se destaca por conter uma estrutura de classificação dos sistemas de IA de acordo com o nível de risco, com medidas de controle e salvaguardas específicas para cada classe.

Sistemas de IA que apresentam risco excessivo à sociedade e aos direitos fundamentais, como aqueles que avaliam traços de personalidade ou preveem a probabilidade de cometimento de crimes com base em comportamento passado, têm seu desenvolvimento e uso vedados. Para sistemas de alto risco, o PL prevê o direito à explicação sobre a decisão automatizada, bem como o direito de contestar e de solicitar a revisão humana dessas decisões.

O PL prevê ainda a governança de sistemas de IA de alto risco, detalhando medidas como a necessidade de documentação do sistema e dos testes, o estabelecimento de mecanismos que permitam apurar potenciais resultados discriminatórios, a implementação de medidas de mitigação de riscos, a adoção de medidas para mitigar e prevenir vieses discriminatórios e a disponibilização de informações que permitam a interpretação dos resultados e o funcionamento de sistemas de IA, de forma que sejam usados de forma ética e responsável.

Para o a utilização de sistemas de IA de alto risco no setor público, o PL acrescenta outras medidas de governança, como a garantia de acesso aos bancos de dados, a portabilidade de dados, a garantia facilitada ao cidadão do direito à explicação e à revisão humana de decisões realizadas pelos sistemas, e a publicização das avaliações dos sistemas de alto risco.

Além disso, o PL estabelece a obrigatoriedade de uma avaliação de impacto algorítmico[15], com medidas preventivas, mitigadoras e de reversão de impactos negativos sobre os direitos fundamentais.

Embora não regule a IA especificamente na área de segurança pública, o PL contribuirá para o uso consciente e controlado de tais sistemas no âmbito da segurança pública, caso seja aprovado.

[15] Conforme dispõe o Inciso XVI, do art. 4º, do PL nº 2.338/2023 , "Avaliação de impacto algorítmico: análise do impacto sobre os direitos fundamentais, apresentando medidas preventivas, mitigadoras e de reversão dos impactos negativos, bem como medidas potencializadoras dos impactos positivos de um sistema de IA".

SGAS Quadra 607 - Módulo 49
Via L2 Sul, Brasília - DF
CEP 70.200-670

(61) 3535-6565

3.2 – RESOLUÇÃO CNJ Nº 615/2025

A Resolução nº 615/2025, do Conselho Nacional de Justiça – CNJ –, estabelece as diretrizes para o desenvolvimento, utilização e governança de soluções que utilizam IA no Poder Judiciário. A norma garante que a aplicação da IA esteja alinhada aos direitos humanos, valores democráticos e à centralidade da pessoa humana, promovendo a igualdade e a justiça decisória. À semelhança do disposto no PL nº 2.338/2023, um pilar central da Resolução é a supervisão humana obrigatória em todas as etapas do ciclo de vida das soluções de IA, assegurando que os sistemas jamais restrinjam ou substituam a autoridade e a decisão final dos magistrados e usuários internos dos sistemas (CNJ, 2025).

Para garantir o uso adequado e ético, a Resolução impõe requisitos de transparência e auditabilidade, com indicadores e relatórios públicos. Soluções classificadas como de alto risco exigem uma avaliação contínua de impacto algorítmico. Além disso, a norma foca na proteção de dados, exigindo curadoria de dados, anonimização e a adoção de práticas de privacidade e proteção de dados pessoais, alinhadas à Lei nº 13.709/2018 – Lei Geral de Proteção de Dados.

A norma previne a discriminação algorítmica ao exigir a implementação de medidas para evitar vieses discriminatórios ilegais ou abusivos. Se um viés for detectado, são exigidas medidas corretivas, como a suspensão, correção ou eliminação definitiva da solução, dada a gravidade dos impactos sobre os direitos fundamentais. Há também proibições explícitas, como o uso de IA para valorar traços de personalidade ou classificar pessoas quanto à previsão de cometimento de crimes ou à probabilidade de reiteração delitiva.

Entre as medidas preventivas para a não ocorrência de vieses, a norma determina que os sistemas de IA devem ter registros de explicações sobre os passos que conduzem ao resultado da tomada de decisão. Além disso, orienta a composição das equipes de pesquisa, desenvolvimento e implantação dos sistemas de IA, de forma que se busque a diversidade de forma ampla, incluindo gênero, etnia, pessoas com deficiência e formação em diferentes áreas.

Ademais, a norma traz regras de prestação de contas, definindo que os modelos de IA deverão assegurar total transparência, compreendendo nomes dos responsáveis pela execução das ações, custos envolvidos na pesquisa e no desenvolvimento do sistema, a existência de ações de colaboração e de cooperação, informação se os resultados pretendidos foram efetivamente alcançados, a demonstração da publicidade quanto ao serviço oferecido, técnicas utilizadas no desenvolvimento e a existência de riscos de erros, contribuindo assim para o impacto positivo aos usuários finais e à sociedade.

SGAS Quadra 607 - Módulo 49
Via L2 Sul, Brasília - DF
CEP 70.200-670

(61) 3535-6565

3.3 – A ESTRATÉGIA E O PLANO BRASILEIRO DE INTELIGÊNCIA ARTIFICIAL

A Estratégia Brasileira de Inteligência Artificial – EBIA –, publicada em 2021, é um documento orientador para o desenvolvimento e uso da IA no Brasil, tanto para o setor público quanto para o privado, estimulando a pesquisa, a inovação e o desenvolvimento de sistemas, de forma consciente e ética (BRASIL, 2021a).

Na EBIA há um capítulo específico para tratar do tema de segurança pública. Conforme levantamento realizado, 75 países já empregam ativamente a IA em tecnologias de vigilância e de segurança pública, voltadas para: otimização do monitoramento e da gestão urbana; reconhecimento facial para a identificação de pessoas em investigação e prevenção de crimes; e para o policiamento inteligente, por meio de análise de dados para prever padrões de criminalidade para subsidiar a atuação das forças de segurança (BRASIL, 2021a, p. 45).

Elemento recorrente no texto da EBIA é a valorização da dignidade humana e do bem-estar humano, que devem estar presentes desde a concepção dos sistemas de IA até a verificação de seus efeitos na vida dos cidadãos – *ethics by design* (BRASIL, 2021a).

A EBIA também aborda a governança de IA, que inclui estratégias como a implantação de comitês de ética, o gerenciamento de riscos, a transparência e a *accountability* – prestação de contas com responsabilização.

Em consonância à EBIA, foi publicado o Plano Brasileiro de Inteligência Artificial – PBIA –, que visa a "desenvolver e implementar tecnologias de IA que impulsionem o progresso econômico e tecnológico do País, atendendo simultaneamente às necessidades reais da população brasileira e respeitando nossa diversidade e valores culturais" (BRASIL, 2025f, p. 14).

Com investimento previsto de R$ 23 bilhões até 2028, o PBIA propõe posicionar o Brasil na vanguarda do desenvolvimento e da aplicação responsável da IA, fomentando a melhoria da qualidade de vida dos brasileiros e a inclusão social, além de disponibilizar soluções concretas em áreas prioritárias como saúde e educação (BRASIL, 2025f, p. 9).

O plano de ação do PBIA está estruturado em cinco eixos[16], detalhando ações, metas, prazos e custos estimados. Especificamente no eixo "IA para Melhoria do Serviço Público, visando eficiência e inovação governamental", é estabelecido o prazo de cinco anos para que 70% dos órgãos federais e 50% dos estaduais implementem soluções de IA com o objetivo de melhorar a eficiência e qualidade dos serviços (BRASIL, 2025f, p. 39). Tal iniciativa possui

[16] Infraestrutura e desenvolvimento de IA; difusão, formação e capacitação em IA; IA para melhoria do serviço público; IA para inovação empresarial; e apoio ao processo regulatório e de governança da IA.

potencial para impulsionar a aplicação de IA também no campo da segurança pública.

Assim, tanto a EBIA quanto a PBIA são importantes instrumentos para os agentes públicos e privados no desenvolvimento e na utilização da IA de forma ética e responsável, promovendo inovação, desenvolvimento econômico e aprimoramento de serviços públicos.

3.4 – PORTARIA MJSP Nº 961/2025

Recentemente, o MJSP publicou a Portaria MJSP nº 961, de 24 de junho de 2025[17], que estabelece diretrizes para o uso de soluções de TI aplicadas às atividades de investigação criminal e inteligência de segurança pública, dirigida a todos os órgãos de segurança pública federais, municipais, estaduais e distritais no escopo dos projetos e ações de segurança pública custeados por recursos federais (BRASIL, 2025e).

A Portaria reforça o respeito a direitos e garantias fundamentais, à proteção de dados pessoais, ao devido processo legal, à transparência, à responsabilização e à prestação de contas.

Entre os objetivos estabelecidos na norma estão a instituição de mecanismos de avaliação e de mitigação de riscos, e a adoção de mecanismos de transparência, auditabilidade, responsabilização e prestação de contas.

Embora represente um avanço na sistematização de procedimentos e orientações aos órgãos de segurança pública quanto ao uso de sistemas de TI, a portaria não aborda especificamente a transparência algorítmica.

A norma estabelece, ainda, que a utilização de soluções de tecnologia aplicadas às atividades de investigação criminal e inteligência de segurança pública "deve respeitar direitos fundamentais e ser limitada ao estritamente necessário para alcançar finalidades compatíveis e circunscritas às competências e atribuições dos órgãos de segurança pública", sendo vedado o seu uso indiscriminado, e sem objetivo certo ou declarado.

Especificamente quanto a soluções de inteligência artificial, a norma permite seu uso em atividades de segurança pública, desde que seus resultados não causem à lesão à vida nem à integridade física das pessoas. Quando houver risco a direitos fundamentais, os agentes de segurança pública deverão revisar o resultado da inferência algorítmica.

Entre as obrigações dos órgãos gestores das soluções de TI, estão a de garantir o uso correto, ético e responsável das soluções, a promoção de capacitações aos usuários, a adoção

[17] Mais informações em: https://www.in.gov.br/en/web/dou/-/portaria-mjsp-n-961-de-24-de-junho-de-2025-638661609. Acesso em: 03 ago. 2025.

de medidas para coibir o uso indevido das soluções, a realização periódica de auditorias e o monitoramento da eficácia das medidas de segurança estabelecidas na norma.

Percebe-se, portanto, considerável evolução na governança tecnológica em segurança pública quanto à utilização de ferramentas de IA, mas ainda há espaço para aprimoramentos, a exemplo do estabelecimento de orientações para a redução da opacidade dos algoritmos.

4 – SISTEMAS DE SEGURANÇA PÚBLICA E INTELIGÊNCIA ARTIFICIAL NO MJSP

Apesar da ausência de um marco regulatório sobre IA, o Brasil possui diversas iniciativas que visam direcionar a sua aplicação para um uso mais consciente e ético da tecnologia, de forma que não lesionem direitos fundamentais, inclusive no âmbito da Administração Pública.

Na temática de segurança pública, evidenciou-se que diversas soluções de IA já são utilizadas pelo mundo, principalmente no reconhecimento facial, em análises preditivas que auxiliam em investigações criminais e na avaliação de propensão à reincidência criminal.

Para atuar na vanguarda tecnológica em segurança pública no âmbito federal, o Ministério da Justiça e Segurança Pública se destaca, dedicando-se ao desenvolvimento e uso de sistemas que utilizam IA para apoiar suas atividades.

Com base nas informações e notícias divulgadas no portal do ministério, os seguintes sistemas, *softwares* e programas utilizam IA ou *Big Data* foram identificados:

Sinesp Big Data – sistema que utiliza IA para prevenir a criminalidade no Brasil, como assaltos, homicídios, monitoramento de veículos roubados e combate ao tráfico em regiões de fronteira. O investimento inicial em infraestrutura digital da solução foi de R$ 32 milhões. Na versão inicial, a solução foi disponibilizada em 2019 a seis estados, com a expectativa de implantação integral em todos os estados da federação, de forma que a solução armazene dados em larga escala, a partir do registro de informações de todos os órgãos da Administração Pública estadual e federal. Com base nos dados registrados, o sistema gera indicadores que auxiliam os órgãos na elaboração de políticas públicas contra a criminalidade nas regiões mapeadas (BRASIL, 2019). Além disso, o sistema registra a geolocalização dos crimes ocorridos, enquanto a IA realiza a predição de ocorrências e propõe rotas para o policiamento ostensivo a partir dos locais e horários dos delitos (BRASIL, 2020a).

Sinesp Agente de Campo – ferramenta digital para acesso policial imediato a

mandados de prisão e para busca nacional de veículos roubados, sendo parte do projeto Sinesp Big Data. Inicialmente foi disponibilizado para três estados, com projeção para disseminação em todo o país (BRASIL, 2020b).

Bancos de Perfis Genéticos – BPG – têm a finalidade de manter, compartilhar e comparar perfis genéticos para auxiliar na elucidação de crimes, verificação de reincidências e na instrução processual. Os perfis genéticos armazenados são confrontados em busca de coincidências que permitam associar suspeitos a locais de crime, incluindo a comparação de perfis genéticos de indivíduos já cadastrados criminalmente, condenados por crimes dolosos praticados com violência grave contra a pessoa, por crimes contra a vida, contra a liberdade sexual ou por crimes sexuais contra vulneráveis (BRASIL, 2021b).

Para ampliar a busca por informações sobre sistemas do MJSP que utilizam IA e *Big Data*, recorreu-se também ao Plano de Dados Abertos do MJSP e ao Portal de Dados Abertos.

Em atenção ao princípio da publicidade e em atendimento à Lei de Acesso à Informação – Lei nº 12.527, de 18 de novembro de 2011 –, os órgãos e entidades públicas têm o dever de publicar os dados e as informações públicas, independentemente de requerimento. Para a operacionalização dessa divulgação, eles devem elaborar um Plano de Dados Abertos, que conterá as orientações para as ações de implementação da abertura de dados, observando-se padrões mínimos de qualidade, de forma a facilitar o entendimento e a reutilização das informações (BRASIL, 2016). Devem também disponibilizar os conjuntos de dados na Internet, por meio do Portal de Dados Abertos, no endereço eletrônico: https://dados.gov.br.

Após análise do Plano de Dados Abertos do MJSP[18], não foram encontradas iniciativas de abertura de dados relacionadas aos sistemas que utilizam IA ou *Big Data*. Foi encontrado apenas um registro de conjunto de dados com indicativo de relacionamento com o ecossistema do Sinesp Big Data: Dados Ocorrências Criminais – Sinesp (BRASIL, 2025c).

Por meio de acesso ao Portal de Dados Abertos[19], identificou-se que o referido conjunto de dados se tratava de dados consolidados de outras soluções, como o SinespJC e o Sinesp Integração, e não dos sistemas que utilizam IA ou *Big Data*.

Como um último recurso para a busca de informações sobre transparência algorítmica nas soluções do MJSP que utilizam IA e *Big Data*, recorreu-se a pedido de informação em 04

[18] Mais informações disponíveis em: https://www.gov.br/mj/pt-br/acesso-a-informacao/dados-abertos/historico-de-planos-de-dados-abertos/pda-2024-2026.pdf/view. Acesso em: 04 ago. 2025.

[19] Mais informações disponíveis em: https://dados.gov.br/dados/conjuntos-dados/sistema-nacional-de-estatisticas-de-seguranca-publica. Acesso em: 04 ago. 2025.

jun. 2025 junto ao órgão, com fundamento no inciso XXXIII[20], art. 5º, da Constituição Federal, e no art. 10[21] da Lei de Acesso à Informação, sendo feitas as seguintes perguntas e pedidos:

> 1 – Além do Sinesp Big Data, Sinesp Geo Inteligência, Sinesp Tempo Real, Sinesp Busca e Banco Nacional de Perfis Genéticos (BNPG), quais outros sistemas de informação finalísticos do MJSP empregam tecnologias de inteligência artificial (IA) ou *Big Data* em suas operações?
>
> 2 – Quais desses sistemas possuem conjuntos de dados abertos publicados no Portal de Dados Abertos? Quais os *links* para *download*?
>
> 3 – Para os sistemas de informação do MJSP que empregam tecnologias de inteligência artificial (IA) ou *Big Data*, incluindo os listados na pergunta 1, existem estudos ou relatórios que avaliam o grau de assertividade, imprecisão, margem de erro ou ocorrência de falsos positivos no processamento de dados? As fontes de erros e incerteza dos algoritmos são identificadas, registradas e comparadas? Se sim, quais os *links* para acesso a esses documentos?
>
> 4 – O MJSP mantém registros ou informações sobre ocorrências de vieses, discriminações ou injustiças resultantes da aplicação de algoritmos de inteligência artificial (IA) e *Big Data* em seus sistemas? Se sim, esses registros ou informações são divulgados? Quais os *links* para acesso/*download*?
>
> 5 – No âmbito do MJSP, existem normativos, procedimentos ou políticas que abordem a reparação ou indenização a indivíduos que comprovem terem sido prejudicados por vieses ou discriminações decorrentes da aplicação de algoritmos de inteligência artificial (IA) e *Big Data*? Se sim, quais os *links* para acesso/*download* desses normativos, procedimentos ou políticas?
>
> 6 – Existe algum procedimento formal ou política no MJSP que garanta que os cidadãos indiciados, suspeitos ou apenados, cujas decisões tenham sido apoiadas por resultados da aplicação de sistemas de inteligência artificial (IA) ou *Big Data* (incluindo os sistemas listados na pergunta 1), sejam informados de que a ação decorreu ou foi subsidiada por tais sistemas? Se sim, poderiam fornecer um exemplo de notificação, termo ou documento que formalize essa informação?
>
> 7 – Para os algoritmos que subsidiam processos decisórios com o uso de inteligência artificial (IA) ou *Big Data* nos sistemas de informação do MJSP (incluindo os listados na pergunta 1), existem documentos que detalham a sua estrutura, funcionamento ou que indiquem a possibilidade de ajuste e parametrização pela equipe técnica do MJSP? Se sim, quais os *links* para acesso a esses documentos?
>
> 8 – Os algoritmos dos sistemas do MJSP que utilizam IA ou *Big Data* que processam dados dos cidadãos são divulgados? Se sim, quais são os *links*?
>
> 9 – Existe regulamento ou política pública no âmbito do MJSP que vise combater a discriminação algorítmica das decisões (ou apoio a decisões) automatizadas de sistemas que utilizam IA ou *Big Data*? Se sim, quais os *links* para *download*?

[20] Inciso XXXIII, art. 5º, da Constituição Federal: "Todos têm direito a receber dos órgãos públicos informações de seu interesse particular, ou de interesse coletivo ou geral, que serão prestadas no prazo da lei, sob pena de responsabilidade, ressalvadas aquelas cujo sigilo seja imprescindível à segurança da sociedade e do Estado".

[21] Art. 10, da Lei de Acesso à Informação: "Qualquer interessado poderá apresentar pedido de acesso a informações aos órgãos e entidades referidos no art. 1º desta Lei, por qualquer meio legítimo, devendo o pedido conter a identificação do requerente e a especificação da informação requerida".

> 10 – Existe algum guia ou norma de uso responsável de ferramentas de IA e *Big Data* no âmbito do MJSP? Se sim, poderiam disponibilizá-lo?
>
> 11 – Para os sistemas de informação que empregam inteligência artificial (IA) ou *Big Data* no MJSP (incluindo os listados na pergunta 1), há documentos, relatórios de auditoria, pareceres ou outros registros que atestem a observância e aplicação do guia ou norma de uso responsável de ferramentas de IA e *Big Data* (mencionado na pergunta 10) no tratamento de dados? Se sim, quais os *links* para acesso a esses documentos?

Em resposta apresentada em 1º jul. 2025, o MJSP informou que não há solução de IA em ambiente de produção que se enquadre nos objetivos descritos no pedido.

Quanto à questão 6, relacionada ao princípio da explicabilidade, o MJSP informou que não aplica recursos automatizados de decisão nem realiza ações com identificação individualizada de cidadãos, que os dados utilizados são exclusivamente para a geração de indicadores estatísticos quantitativos e agregados, sem qualquer identificação individualizada de cidadãos.

Sobre a questão 10, o ministério informou que segue as recomendações e diretrizes estabelecidas na cartilha elaborada pelo Ministério da Gestão e da Inovação em Serviço Público "IA Generativa no Serviço Público - definições, usos e boas práticas"[22], e a recém elaborada "Cartilha de Boas Práticas no Uso de Inteligência Artificial Generativa do MJSP", que contém orientações aos servidores sobre o uso seguro e ético de ferramentas externas ou comerciais de Inteligência Artificial Generativa.

Em que pese tais cartilhas trazerem valiosas contribuições para o desenvolvimento e uso da IA, elas tratam apenas da IA generativa (BRASIL, 2025b; BRASIL, 2025d), uma subclasse da IA, não abarcando de forma geral os sistemas de IA e *Big Data* já apresentados.

De toda forma, a cartilha elaborada pelo MJSP apresenta preocupações e valores gerais que podem ser aproveitados na operação de sistemas de IA que não são generativas, dada a natureza intrínseca dos modelos de IA que utilizam dados históricos para treinamento e aprendizado, a exemplo do respeito à autonomia humana, prevenção de danos, não discriminação, explicabilidade, transparência, sustentabilidade, respeito aos direitos humanos e democráticos, e segurança (BRASIL, 2025d).

Apesar de haver notícias no portal do MJSP sobre soluções implantadas que utilizam IA e *Big Data*, o ministério não apresentou respostas para as demais questões, atinentes a pontos cruciais para o adequado desenvolvimento e utilização de sistemas de IA, tais como: validação

[22] Mais informações disponíveis em: https://www.gov.br/governodigital/pt-br/infraestrutura-nacional-de-dados/inteligencia-artificial-1/ia-generativa-no-servico-publico.pdf. Acesso em: 04 ago. 2025.

das informações produzidas pelos sistemas, transparência sobre o funcionamento dos algoritmos, registro de ocorrência de vieses e discriminação, análise da acurácia das decisões automatizadas, responsabilidade e reparação de danos às pessoas impactadas pelos vieses e falsos positivos, e direito de acesso à informação sobre decisões automatizadas apoiadas pelas ferramentas. Essa ausência de informações evidencia a necessidade de aprimoramento da governança de sistemas de IA da organização.

5 – FORMAS DE MITIGAR A DISCRIMINAÇÃO ALGORÍTMICA

Conforme apresentado, apesar de haver notícias sobre o desenvolvimento e a utilização de sistema de IA pelo MJSP para aplicação em segurança pública, não há uma transparência mínima sobre o funcionamento dos algoritmos, que poderia incluir regras básicas, passos, natureza dos dados e informações sobre validação do treinamento e análise de riscos.

A ausência de informações é compreensível caso os sistemas de IA ainda estivessem em um estágio incipiente de desenvolvimento e não estejam em ambiente de produção. Contudo, as diretrizes da EBIA e demais normas analisadas indicam a importância de considerar a ética e as boas práticas desde a concepção dos sistemas, com vistas a valorizar a dignidade e o bem-estar humano, além de mitigar os vieses e as ocorrências de discriminação (BRASIL, 2021a).

Como a maioria dos modelos de IA dependem de treinamento que utilizam bases de dados históricos, eles podem intrinsicamente gerar resultados enviesados e discriminatórios (JUNQUILHO e DIAS, 2024). Isso reforça a importância da supervisão humana e da avaliação criteriosa, especialmente em órgãos públicos que processam informações sensíveis e de alto impacto social (BRASIL, 2025b).

Sabendo-se desses riscos, a EBIA orienta as organizações e desenvolvedores de sistemas de IA a considerar os princípios nela apresentados[23], e verificar periodicamente se estão sendo respeitados (BRASIL, 2021a).

Além da EBIA, o setor público brasileiro conta com outros normativos que orientam o desenvolvimento e o uso responsável e ético de sistemas de IA, conforme apresentado na Seção 3, cujas principais diretrizes e práticas são apresentadas a seguir:

Instituição de comitês de ética: promovem a responsabilidade sobre o uso adequado da IA nas organizações, promovem tomadas de decisões responsáveis e asseguram que novas

[23] "(i) Crescimento inclusivo, desenvolvimento sustentável e o bem-estar; (ii) valores centrados no ser humano e na equidade; (iii) transparência e explicabilidade; (iv) robustez, segurança e proteção e; (v) a responsabilização ou a prestação de contas – *accountability*" (BRASIL, 2021a, p. 17).

utilizações de dados respeitem os valores corporativos e sociais (BRASIL, 2021a);

Realização de análises de riscos: podem ser consubstanciados por meio da elaboração de relatórios de impacto, que registram como as organizações avaliam questões de justiça e direitos humanos na implantação de novas tecnologias de IA, além de medidas para a sua implementação (BRASIL, 2021a; COELHO, 2024);

Promoção da transparência: consiste na adoção de metodologias transparentes e auditáveis para o desenvolvimento dos sistemas de IA, incluindo fontes de dados, procedimentos e documentação dos projetos (BRASIL, 2021a), além da utilização preferencial de *software* de código aberto, que facilita a integração e a interoperabilidade do modelo com outros sistemas, além de permitir uma maior cooperação com outros segmentos e áreas do setor público e da sociedade civil (CNJ, 2025);

Explicabilidade: define que as decisões tomadas por sistemas automatizados sejam passíveis de explicação e de interpretação, de como os dados das pessoas foram utilizados na decisão automatizada, mesmo nos casos de sistemas mais fechados (BRASIL, 2021a; (BRASIL, 2025a). Nas palavras de Mendes e Mattiuzzo (2019, p. 56), a explicabilidade se configura como uma "descrição, compreensível por humanos, do processo por meio do qual aquele que toma a decisão, ao utilizar um certo grupo de *inputs*, atinge uma dada conclusão".

Accountability: traduzida como responsabilidade com prestação de contas. Impõe que sejam estabelecidas estruturas de governança de IA que possam assegurar a adoção de princípios para IA confiável e implementar mecanismos para sua observância (BRASIL, 2021a). Mendes e Mattiuzzo (2019) vão além, informando que pessoas serão afetadas pelo processo decisório algorítmico, sendo necessário reconhecer a responsabilização e oferecer alternativas a eventual reparação de danos.

Em contribuição, Coelho (2024, p. 6-7) apresenta recomendações aos usuários e desenvolvedores de sistemas de IA em políticas públicas, destacando: a necessidade de reconhecer que os dados utilizados nos algoritmos são frequentemente incompletos e enviesados, o que requer análise crítica contínua sobre os resultados; a importância de marcos regulatórios robustos para garantir que os algoritmos sejam desenvolvidos e utilizados de maneira ética e responsável; que os algoritmos sejam projetados com o objetivo de reduzir, e não de amplificar, os vieses existentes nos dados de treinamento; e que as decisões tomadas por IA em contextos de justiça e segurança devem ser compreensíveis e passíveis de contestação.

Sobre normatização, Mendes e Mattiuzzo (2019) e Silva (2022) ressaltam a necessidade

SGAS Quadra 607 - Módulo 49
Via L2 Sul, Brasília - DF
CEP 70.200-670

(61) 3535-6565

de soluções regulatórias e de políticas públicas para enfrentar o problema da discriminação presente nos processos decisórios de IA. A ausência de transparência sobre os algoritmos levanta sérias preocupações quanto às consequências legais da discriminação algorítmica.

Na mesma linha, Junquilho e Dias (2024, p. 135) reforçam a necessidade de legislações específicas para limitar o uso da IA, que considerem princípios éticos, transparência, privacidade e não discriminação, em alinhamento aos valores fundamentais da sociedade.

Nesse aspecto, já se observa movimentação para a aprovação do Marco Regulatório da Inteligência Artificial no Brasil, por meio do PL nº 2338/2023, que contém elementos fundamentais para o desenvolvimento e uso consciente e ético da IA.

No âmbito da segurança pública em nível federal, o MJSP demostra considerável avanço nas normas e guias para o desenvolvimento e utilização de sistemas de IA.

Conforme tratado por Mendes e Mattiuzzo (2019), não há consenso na literatura de governança algorítmica quanto à importância e utilidade da transparência sobre os algoritmos, havendo uma linha de autores que argumenta pela auditabilidade em vez da transparência, pois nem sempre os códigos dos algoritmos podem ser publicizados.

Além disso, revelar informações demais sobre algoritmos e processos de sistemas de IA pode tanto sobrecarregar as pessoas com dados irrelevantes quanto expor segredos comerciais, industriais e de propriedade intelectual. Portanto, a explicabilidade deve se concentrar em fornecer apenas informações significativas que permitam interpretar o sistema, sem sobrecarregar ou expor detalhes sensíveis. (BRASIL, 2021a).

Em uma abordagem de máxima transparência, os dados e os passos utilizados pelo algoritmo no processo decisório devem estar disponíveis para conhecimento dos usuários, com vistas a identificar os vieses para providências de correções. Essa transparência não estaria restrita apenas a autoridades, mas também ao público em geral, incluindo a abertura das bases de dados, do código-fonte e da modelagem dos sistemas (MENDES e MATTIUZZO, 2019).

Para os sistemas de IA e *Big Data* do MJSP pesquisados, apenas os Bancos de Perfis Genéticos possuem alguma transparência sobre a sua operação, mas de forma estática, registrada no Manual de Procedimentos Operacionais da Rede Integrada de Bancos de Perfis Genéticos[24], apresentando como são feitas as análises estatísticas e interpretação dos resultados, as tarefas de revisão, a confirmação e a classificação dos resultados (BRASIL, 2024). De toda

[24] Mais informações disponíveis em: https://www.gov.br/mj/pt-br/assuntos/sua-seguranca/seguranca-publica/ribpg/manual/manual-de-procedimentos-operacionais-da-ribpg-versao-6-1-aprovado-pela-resolucao-no-10-de-6-de-marco-de-2025-do-comite-gestor-da-ribpg.pdf . Acesso em 03 ago. 2025.

forma, não foram encontradas informações sobre o registro de ocorrência de vieses e discriminação, e nem de casos de responsabilização e reparação de danos a pessoas impactadas pelos vieses e falsos positivos.

Quanto às demais soluções pesquisadas – Sinesp Big Data e Sinesp Agente de Campo –, não se identificou nenhuma iniciativa de transparência de dados, dos algoritmos e dos resultados de sua aplicação. A busca incluiu pesquisa no portal do ministério, no Portal de Dados Abertos e solicitação de acesso à informação, prevista na Lei de Acesso à Informação.

Uma hipótese de causa da não divulgação das informações dos sistemas de IA seria o comprometimento da segurança das operações policiais. Com o conhecimento prévio pelos criminosos sobre quais são os parâmetros e qual é a lógica dos algoritmos, poderiam prever locais de patrulhamento ou identificar padrões para adaptar suas ações e burlar o sistema.

Nas situações em que a transparência total possa comprometer a segurança de operações policiais, uma abordagem de transparência intermediária poderia ser uma alternativa mais factível, em que um conjunto mínimo de informações[25] sobre o funcionamento dos algoritmos seria divulgado, cuja definição poderia ser analisada por um comitê de ética, conforme defendido na EBIA.

Aliado a isso, medidas com intervenção humana podem ser utilizadas para mitigar o risco de discriminação algorítmica, como a garantia de supervisão humana sobre o processo decisório do algoritmo, a realização de auditorias independentes[26] e o treinamento das equipes que desenvolvem, validam e utilizam os sistemas de IA (MENDES e MATTIUZZO, 2019; JUNQUILHO e DIAS, 2024; BRASIL, 2025a; BRASIL, 2025b; BRASIL, 2025e; CNJ, 2025).

6 – CONSIDERAÇÕES FINAIS

A presente pesquisa buscou analisar os impactos da discriminação algorítmica no âmbito da segurança pública no Brasil, a partir da identificação dos sistemas de IA e *Big Data* utilizados pelo MJSP e das medidas implementadas para mitigar os problemas. Para tanto, empregou-se uma metodologia que combinou pesquisa bibliográfica, busca por informações no portal do ministério, no Portal de Dados Abertos e um pedido de acesso à informação. Os

[25] A exemplo do objetivo do sistema, quais tipos de dados são utilizados no treinamento do sistema, se houve auditoria independente sobre o sistema, os riscos envolvidos, informações sobre a avaliação de impacto e da acurácia do sistema, o nível de diversidade nas equipes de concepção, desenvolvimento e aplicação do sistema e se há protocolos que garantam a qualidade dos dados (INSTITUTO IGARAPÉ, 2025).

[26] Além de auditorias que podem ser realizadas órgãos públicos formalmente constituídos, como a Controladoria-Geral da União e o Tribunal de Contas da União, a realização de auditorias independentes por empresas especializadas em IA poderiam trazer um reforço quanto ao atendimento das premissas de não discriminação.

resultados deste estudo mostram que, embora o Brasil tenha avançado no estabelecimento de diretrizes e práticas para o uso ético e responsável da IA, inclusive no âmbito do MJSP, ainda há um considerável espaço para aprimoramento, especialmente no que tange à transparência e à responsabilidade no uso dessas tecnologias, além da necessidade de consolidação do marco regulatório da IA no Brasil.

A ascensão da IA e do *Big Data* na segurança pública, embora promissora, traz consigo o risco inerente de perpetuar e amplificar discriminações sociais preexistentes. A pesquisa demonstrou que a principal causa desse fenômeno reside no uso de bases de dados incompletas, enviesadas ou não representativas, o que pode levar a resultados discriminatórios, como os falsos positivos, o perfilamento racial e a discriminação. Além disso, a ausência de transparência nos algoritmos e nos processos de tomada de decisão automatizada gera desconfiança e insegurança, dificultando a identificação de vieses e suas correções.

No âmbito normativo, verificou-se que o Brasil dispõe de importantes instrumentos para o uso ético e responsável da IA, como a Estratégia e o Plano Brasileiro de Inteligência Artificial, a Resolução CNJ nº 615/2025 e o Projeto de Lei nº 2.338/2023, que abordam princípios como a não discriminação, a explicabilidade, a transparência e a *accountability*. A recente Portaria MJSP nº 961/2025 também representa um avanço significativo ao estabelecer diretrizes para o uso de tecnologias na segurança pública. Embora a portaria mencione sistemas de IA, ela não trata da transparência algorítmica. Incluir esse aspecto seria crucial para reduzir a opacidade dos algoritmos, de forma que as pessoas afetadas por decisões automatizadas possam entender como a decisão foi tomada, dando-lhes subsídios para contestar e solicitar a revisão humana.

A despeito da existência de sistemas como o Sinesp Big Data e o Sinesp Agente de Campo, divulgados no portal do ministério como exemplos de sistema de IA e *Big Data* do MJSP, a busca por informações sobre sua operação não obteve êxito quanto à transparência dos dados e dos algoritmos.

Este estudo reitera que a mera existência de normas não garante a sua efetiva aplicação, sendo a falta de transparência algorítmica um obstáculo fundamental para a garantia de direitos fundamentais e a mitigação de danos. Além do aprimoramento das normas, é importante que sejam estabelecidos mecanismos práticos de governança de IA, como a implantação de comitê de ética que trate do tema, incentivo e aplicação de práticas de *accountability*, gestão de riscos com análise de impacto, facilitação do acesso aos dados abertos, além da realização de campanhas educacionais e de conscientização, conforme apresentados na EBIA e no PBIA.

SGAS Quadra 607 - Módulo 49
Via L2 Sul, Brasília - DF
CEP 70.200-670

(61) 3535-6565

Em suma, a solução proposta para mitigar a discriminação algorítmica em segurança pública alinha-se às boas práticas de governança já discutidas, com um foco especial na transparência. De toda forma, reconhece-se que a total publicização de algoritmos pode comprometer a eficácia das operações policiais e o sigilo de segredos comerciais, industriais e de propriedade intelectual. Assim, a proposta sugere a adoção de uma abordagem intermediária, que consiste na divulgação de um conjunto mínimo de informações sobre o funcionamento dos algoritmos, a ser definido ou validado por um comitê de ética.

Ademais, medidas com intervenção humana podem ser utilizadas para mitigar o risco de discriminação algorítmica, como a garantia de supervisão humana sobre o processo decisório do algoritmo, a realização de auditorias independentes, o treinamento das equipes que desenvolvem, validam e utilizam os sistemas de IA, além do arranjo multidisciplinar e representativo de pessoas que formam as equipes de desenvolvimento, para que se reduzam os vieses desde a fase de concepção – *ethics by design*.

Em última análise, a transparência algorítmica em sistemas de IA para a segurança pública é um pilar crucial para assegurar que a tecnologia sirva à justiça e assegure os direitos fundamentais para todos.

REFERÊNCIAS

BBC. **Google pede desculpas por erro racista no aplicativo Fotos.** 2015. Disponível em: https://www.bbc.com/news/technology-33347866.Acesso em: 23 out. 2025.

BRASIL. **Constituição da República Federativa do Brasil, 1988.** Disponível em: https://www.planalto.gov.br/ccivil_03/constituicao/constituicaocompilado.htm. Acesso em: 29 jul. 2025.

BRASIL. **Decreto nº 8.777, de 11 de maio de 2016.** Institui a Política de Dados Abertos do Poder Executivo federal. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2016/decreto/D8777.htm. Acesso em: 04 ago. 2025.

BRASIL. **Lei nº 13.675, de 11 de junho de 2018.** Disciplina a organização e o funcionamento dos órgãos responsáveis pela segurança pública [...], institui o Sistema Único de Segurança Pública [...]. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13675.htm. Acesso em: 29 jul. 2025.

BRASIL. Ministério da Justiça e Segurança Pública. **MJSP realiza primeiras entregas do Big Data e Inteligência Artificial**. 2019. Disponível em: https://www.gov.br/mj/pt-br/centrais-de-conteudo/videos/mjsp-realiza-primeiras-entregas-do-big-data-e-inteligencia-artificial. Acesso em: 29 jul. 2025.

BRASIL. Universidade Federal do Ceará. **SINESP Big Data: Centro de Referência em Inteligência Artificial já desenvolve primeiro projeto, e se prepara para outro**. 2020a. Disponível em https://www.ufc.br/noticias/noticias-de-2020/15087-sinesp-big-data-centro-de-

referencia-em-inteligencia-artificial-ja-desenvolve-primeiro-projeto-e%E2%80%A6. Acesso em: 03 ago. 2025.

BRASIL. Ministério da Justiça e Segurança Pública. **Ministério da Justiça e Segurança Pública lança aplicativo Sinesp Agente de Campo**. 2020b. Disponível em: https://www.gov.br/mj/pt-br/assuntos/noticias/ministerio-da-justica-e-seguranca-publica-lanca-aplicativo-sinesp-agente-de-campo. Acesso em: 03 ago. 2025.

BRASIL. Ministério da Ciência, Tecnologia e Inovação. **Estratégia Brasileira de Inteligência Artificial – EBIA**. 2021a. Disponível em: https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/arquivosinteligenciaartificial/ebia-diagramacao_4-979_2021.pdf. Acesso em: 10 ago. 2025.

BRASIL. Ministério da Justiça e Segurança Pública. **A Rede Integrada de Bancos de Perfis Genéticos (RIBPG)** [...]. 2021b. Disponível em: https://www.gov.br/mj/pt-br/assuntos/sua-seguranca/seguranca-publica/ribpg/institucional. Acesso em: 03 ago. 2025.

BRASIL. Ministério da Justiça e Segurança Pública. **Manual de Procedimentos Operacionais da Rede Integrada de Bancos de Perfis Genéticos.** Versão 6.1, 20 dez. 2024. Disponível em: https://www.gov.br/mj/pt-br/assuntos/sua-seguranca/seguranca-publica/ribpg/manual/manual-de-procedimentos-operacionais-da-ribpg-versao-6-1-aprovado-pela-resolucao-no-10-de-6-de-marco-de-2025-do-comite-gestor-da-ribpg.pdf. Acesso em: 03 ago. 2025.

BRASIL. Câmara dos Deputados. **Projeto de Lei nº 2338/2023.** Dispõe sobre o desenvolvimento, o fomento e o uso ético e responsável da inteligência artificial com base na centralidade da pessoa humana. Última ação legislativa: 23 out. 2025a. Disponível em: https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2487262. Acesso em: 26 out. 2025.

BRASIL. Ministério da Gestão e da Inovação em Serviços Públicos. **IA Generativa No Serviço Público Definições, usos e boas práticas.** 2025b. Disponível em: https://www.gov.br/governodigital/pt-br/infraestrutura-nacional-de-dados/inteligencia-artificial-1/ia-generativa-no-servico-publico.pdf. Acesso em: 10 ago. 2025.

BRASIL. Ministério da Justiça e Segurança Pública. **Plano de Dados Abertos do MJSP – 2024-2026**. 2025c. Disponível em: https://www.gov.br/mj/pt-br/acesso-a-informacao/dados-abertos/historico-de-planos-de-dados-abertos/pda-2024-2026.pdf/view. Acesso: 04 ago. 2025.

BRASIL. Ministério da Justiça e Segurança Pública. **Cartilha de Boas Práticas no Uso de Inteligência Artificial Generativa**. Acervo particular do autor. Acesso em: 01 jul. 2025d.

BRASIL. Ministério da Justiça e Segurança Pública. **Portaria MJSP nº 961, de 24 de junho de 2025.** Estabelece diretrizes sobre uso de soluções de tecnologia da informação aplicadas às atividades de investigação criminal e inteligência de segurança pública. 2025e. Disponível em: https://www.in.gov.br/en/web/dou/-/portaria-mjsp-n-961-de-24-de-junho-de-2025-638661609. Acesso em: 03 ago. 2025.

BRASIL. Ministério da Ciência, Tecnologia e Inovações. **Plano Brasileiro de Inteligência Artificial (PBIA) 2024-2028,** 2025f. Disponível em: https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/noticias/2025/06/plano-brasileiro-de-inteligencia-artificial-pbia-_vf.pdf. Acesso em: 25 out. 2025.

CNJ. Conselho Nacional de Justiça. **Resolução nº 615, de 11 de março de 2025.** Estabelece diretrizes para o desenvolvimento, utilização e governança de soluções desenvolvidas com recursos de inteligência artificial no Poder Judiciário. Disponível em:

https://atos.cnj.jus.br/files/original155530202503146 7d4517244566.pdf. Acesso: 12 out. 2025.

COELHO, Danilo S. C. **Inteligência Artificial em Justiça e Segurança Pública: Exemplos e Recomendações para Políticas Públicas.** Rio de Janeiro: Instituto de Pesquisa Econômica e Aplicada. 1ª Ed. 2024. Disponível em: https://repositorio.ipea.gov.br/server/api/core/bitstreams/e019a1ad-ec6f-4c62-9468-10aed0572bf1/content. Acesso em: 23 ago. 2025.

DASTIN, Jeffrey. **Amazon scraps secret AI recruiting tool that showed bias against women**. Reuters, 2018. Disponível em: https://www.reuters.com/article/us-amazon-com-jobsautomation-insight-idUSKCN1MK08G. Acesso em: 24 out. 2025.

DOMINGOS, P. **The master algorithm: how the quest for the ultimate learning machine will remake our world**. New York: Basic Books, 2015. 352p.

IBM. International Business Machines Corporation. **O que é Big Data?** Disponível em: https://www.ibm.com/br-pt/think/topics/big-data. Acesso em: 19 ago. 2025.

INSTITUTO IGARAPÉ. **Guia para Uso Responsável, Transparente e Seguro da Inteligência Artificial na Segurança Pública**. Disponível em: https://igarape.org.br/wp-content/uploads/2022/05/Guia-da-Inteligencia-Artificial-na-Seguranca-Publica.pdf. Acesso em: 25 ago. 2025.

JUNQUILHO, T. A.; DIAS, J. M. P. Racismo algorítmico: uma análise sobre os riscos do uso do reconhecimento facial pelos órgãos de segurança pública. In: Bia Barbosa; Laura Tresca; Luana Roncaratti; Mozart Tenório; Renata Mielli; Tanara Lauschner. (Org.). **TIC, Governança da Internet, Gênero, Raça e Diversidade.** 1ª ed. São Paulo: NICBR, 2024, v. 1, p. 125-150.

MENDES, Laura S.; MATTIUZZO, Marcela. **Discriminação Algorítmica: Conceito, Fundamento Legal e Tipologia.** RDU, Porto Alegre, Vol. 16, nº 90, p. 39-64, nov. 2019.

NUNES, P. **Pablo Nunes: racismo algorítmico e segurança pública.** Nexo, 2022 (entrevista). Disponível em: https://pp.nexojornal.com.br/pergunte-a-um-pesquisador/2022/02/02/pablo-nunes-racismo-algoritmico-e-seguranca-publica. Acesso em: 23 out. 2025.

ONU. Organização das Nações Unidas. **Declaração Universal dos Direitos Humanos.** Proclamada em 10 dez. 1948. Disponível em: https://www.unicef.org/brazil/declaracao-universal-dos-direitos-humanos. Acesso em: 18 ago. 2025.

SAINZ, Nilton; GABARDO, Emerson; ONGARATTO, Natália. **Discriminação Algorítmica no Brasil: Uma Análise da Pesquisa Jurídica e suas Perspectivas para a Compreensão do Fenômeno.** Revista Direito Público. Brasília. Vol.21 n. 110, p. 258-289, abr./jun. 2024.

SÃO PAULO. **Smart Sampa.** Disponível em: https://smartsampa.prefeitura.sp.gov.br/. Acesso em: 23 ago. 2025.

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

VEJA. **Decolar é multada em R$ 7,5 milhões por diferenciar preços a consumidor**, 2018. Disponível em: https://veja.abril.com.br/economia/decolar-e-multada-em-r-75-milhoes-por-diferenciar-precos-a-consumidor/. Acesso em 24 out. 2025.

VINCENT, James. **Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day.** The Verge**,** 2016. Disponível em: https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist. Acesso em: 24 out. 2025.