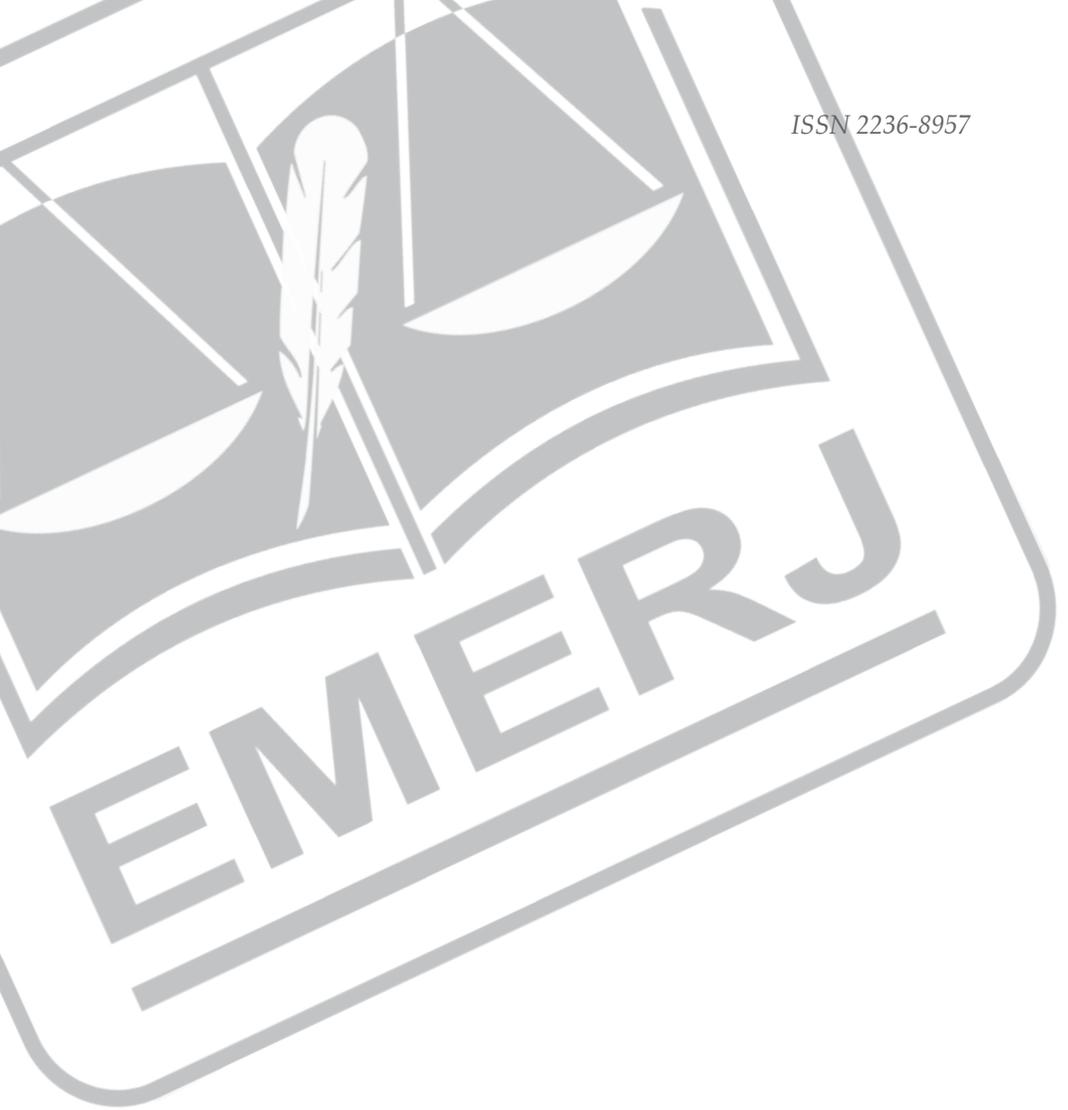


ISSN 2236-8957



Revista da EMERJ

Janeiro/Abril
V. 25 - n. 1 - Ano 2023

Rio de Janeiro

A Inteligência Artificial e o Processo Penal: A Utilização da Técnica na Violação de Direitos

Catiane Steffen

Doutoranda na Pontifícia Universidade Católica do Rio Grande do Sul (Direito). Bacharela em Direito. Pesquisadora e cientista.

SUMÁRIO: 1. Introdução. 2. Desenvolvimento. 2.1 Conceitos e noções iniciais de inteligência artificial. 2.2 A utilização da inteligência artificial na violação de direitos no processo penal. 3. Conclusão. Referências

RESUMO: Este artigo apresenta um estudo introdutório sobre como a inteligência artificial pode reproduzir a discriminação e outras violações de direitos na persecução penal. A metodologia do trabalho se ampara na abordagem dialética, e a técnica de pesquisa, na documentação indireta, especialmente bibliográfica. Como principal conclusão, merece ser destacado que, diferentemente do que alguns autores sustentam – que a tomada de decisão realizada por procedimentos automatizados seria mais objetiva, consistente e neutra –, o que se verifica na prática é que a tomada de decisão orientada por mecanismos que se utilizam de estratégias de emulação do comportamento humano pode resultar na potencialização da violação de direitos. A principal contribuição deste trabalho é demonstrar a necessidade de se reconhecer que o desenvolvimento, a exploração e o uso da inteligência artificial ou de qualquer algoritmo deve acontecer num espaço de comprometimento com o respeito e a promoção dos direitos humanos.

PALAVRAS-CHAVE: Algoritmos – Discriminação – Inteligência Artificial – Processo Penal – Violação de Direitos.

ABSTRACT: This article presents an introductory study on how artificial intelligence may reproduce discrimination and other rights violations in criminal prosecution. The methodology of this work is based on the dialectical approach, and the research technique is indirect documentation, especially a bibliographical review. The main conclusion is that, unlike what some authors claim that decision making performed by automated procedures would be more objective, consistent, and neutral, what actually happens is that decision making guided by mechanisms that use strategies to emulate human behavior may result in the enhancement of rights violation. The main contribution of this work is to demonstrate the need to recognize that the development, exploration, and use of artificial intelligence or of any algorithm must take place in a space of commitment to the respect and promotion of human rights.

KEYWORDS: Algorithms – Artificial Intelligence – Criminal Procedure – Discrimination – Violation of Rights.

RESUMEN: Este artículo presenta un estudio introductorio sobre como la inteligencia artificial puede reproducir la discriminación y otras violaciones de derechos en la persecución penal. La metodología del trabajo se ampara en el abordaje dialéctico y la técnica de pesquisa en la documentación indirecta, especialmente bibliográfica. Como principal conclusión merece ser destacado que diferentemente de lo que algunos autores sustentan que la tomada de decisión realizada por procedimientos automatizados sería más objetiva, consistente y neutra, lo que se verifica en la práctica es que la tomada de decisión orientada por mecanismos que utilizan estrategias de emulación del comportamiento humano puede resultar en la potencialización de la violación de derechos. La principal contribución de este trabajo es demostrar la necesidad de reconocer que el desarrollo, la explotación y el uso de la inteligencia artificial o de cualquier algoritmo debe suceder en un espacio de comprometimiento con el respeto y la promoción de los derechos humanos.

PALABRAS-LLAVE: Algoritmos – Discriminación – Inteligencia Artificial – Proceso Penal – Violación de Derechos.

RÉSUMÉ: Cet article présente une étude introductive sur la façon dont l’intelligence artificielle peut reproduire la discrimination et d’autres violations de droits et de persécution pénale. La méthodologie du travail s’ancre à l’approche dialectique et la technique de recherche dans la documentation indirecte, spécialement bibliographique. Comme conclusion principale il mérite d’être souligné que, différemment de ce que certains auteurs soutiennent à propos de l’utilisation de procédures automatisées pour une prise de décision plus objective, consistante et neutre, ce qui est vérifié dans la pratique est qu’il est possible qu’une prise de décision orientée par des mécanismes qui s’utilisent de stratégies d’émulation du comportement humain, origine l’intensification de la violation de droits. La principale contribution de ce travail est de démontrer le besoin de reconnaître que le développement, l’exploration et l’utilisation de l’intelligence artificielle ou tout algorithme doit se passer dans un espace de compromis par le respect et la promotion des droits humains.

MOTS-CLÉS: Algorithmes – Discrimination – Intelligence Artificielle – Procédure Pénale – Violation de Droits.

1 INTRODUÇÃO

A inteligência artificial é parte de um processo de revolução digital que cada vez mais demonstra protagonismo na sociedade. Alguns autores colocam a inteligência artificial como um dos desenvolvimentos da chamada “quarta revolução industrial” (Schwab, 2016). Enquanto a primeira revolução industrial foi marcada pela invenção da máquina a vapor, a segunda revolução industrial teve como propulsor o advento da eletricidade e da linha de montagem, que permitiu a produção em massa dos produtos. A terceira revolução industrial abriu a revolução digital, na década de 1960, impulsionada pelo desenvolvimento de tecnologias – como a internet – e de elementos de hardware computacional. Por fim, a quarta revolução industrial, conforme Schwab (2016), é caracterizada por três vetores.

O primeiro vetor é a velocidade, pois avança num ritmo exponencial no qual as novas tecnologias dão origem a tecnologias cada vez mais recentes e melhoradas. O segundo vetor é a amplitude e a profundidade, pois combina várias inovações em diversas áreas, alterando a forma como se faz as coisas. Por fim, o terceiro vetor é o impacto sistêmico, pois os seus efeitos atingem não somente uma sociedade local, mas o mundo, seja nos negócios, na política ou nas relações sociais de modo geral.

A inteligência artificial enquanto conceito tem diferentes definições na literatura, que são abordadas a seguir, na Seção 2, especificamente na Subseção 2.1, mas todas convergem numa mesma direção: a da emulação do comportamento humano. Na Seção 3, apresenta-se uma revisão da literatura da área sobre a violação de direitos a partir da aplicação de técnicas como a inteligência artificial, destacando-se aspectos e casos que envolvem o processo penal. Em seguida, são apresentadas as considerações finais.

No mundo fático, as violações de direitos costumam ser percebidas mais facilmente quando são materializadas pela ação ou omissão de agentes humanos do que quando são resultantes de processamento e de decisões produzidas por algoritmos. Isso porque, computacionalmente, as regras e as definições que estruturam a lógica dos programas e que podem determinar os resultados produzidos costumam estar encobertas pela opacidade (ausência de transparência), ou, ainda, pela imprevisibilidade e pela autonomia de alguns sistemas complexos.

A fim de tornar o texto de melhor compreensão para os leitores e de padronizar a escrita à forma como outros autores se expressam no Direito, utiliza-se neste trabalho o termo “algoritmo” de um modo mais amplo, sem o rigor técnico da palavra, englobando também aquilo que na Computação entende-se como “código”. Essa intercambialidade entre os termos é adotada na escrita especificamente pelos motivos citados.

2 DESENVOLVIMENTO

A inteligência artificial é apresentada nesta seção por meio de definições expressas na Subseção 2.1. Essas definições são importantes porque, quando se pensa em inteligência artificial, tem que se entender que a técnica não garante a efetiva inteligência de um algoritmo e não assegura que as decisões e os resultados produzidos serão melhores do que aqueles que um ser humano é capaz de produzir. Em seguida, na Subseção 2.2, apresentam-se estudos que demonstram a utilização da técnica na conformação de cenários de discriminação algorítmica na persecução penal.

2.1 Conceitos e noções iniciais de inteligência artificial

No trabalho de Kaplan e Haenlein (2019), os autores definem a inteligência artificial como a habilidade de um sistema de interpretar corretamente dados externos, aprender a partir desses dados e utilizar esses aprendizados para atingir metas e objetivos específicos por meio de uma adaptação flexível. Os autores Poole e Mackworth (2017) definem inteligência artificial como um campo que estuda a síntese e a análise de agentes computacionais que agem de maneira inteligente. Em Russell e Norvig (2020), os autores definem o conceito de inteligência artificial como o estudo de agentes (inteligentes) que recebem preceitos do ambiente e agem.

Nas últimas duas décadas, houve o desenvolvimento de novas técnicas computacionais. Essas técnicas, combinadas com o aumento da capacidade do hardware disponível e o uso de conjuntos de dados para o treinamento e o desenvolvimento de aprendizado de máquina, permitiram a aplicação de algoritmos matemáticos no campo da inteligência artificial.

Desse modo, viabilizou-se que tarefas humanas mais complexas fossem desenvolvidas e implementadas por meio de algoritmos. A evolução das técnicas computacionais chegou ao ponto de ser possível que um algoritmo, a partir de um conjunto de dados iniciais, consiga estabelecer funções, relações e fazer

previsões de dados sem a necessidade de intervenção humana direta no processo de treinamento.

A expressão “inteligência artificial” tem sido usada com frequência de modo intercambiável com outros termos, como “aprendizagem de máquina”, a qual pode ser supervisionada (que se refere a algoritmos e modelos estatísticos que aprendem com os dados de treinamento fornecidos) ou não supervisionada (quando reconhecem novos padrões que são identificados nos dados processados sem a influência de nenhum conhecimento prévio). Por conseguinte, a aprendizagem de máquina permite reconhecer e inferir padrões, assim como executar tarefas específicas sem a necessidade de instruções explícitas de um operador humano.

Um termo muito utilizado quando se fala em inteligência artificial é “agente”, que é algo ou alguém que age. No entanto, nem todo agente é um agente inteligente. Na definição de Bartneck *et al.* (2020), um agente pode ser definido como inteligente quando exerce ações apropriadas para as suas circunstâncias e para os seus objetivos, é flexível e compreende as mudanças de ambiente e de metas, aprende com a experiência e faz escolhas apropriadas dadas as suas limitações tanto de percepção quanto computacionais.

A inteligência artificial pode ser classificada em vários níveis, que incluem forte (geral) (Strauß, 2018) e fraca (Mántaras, 2019). A inteligência artificial forte é aquela em que o agente emula um comportamento humano de tal maneira que o torna quase indistinguível do comportamento proveniente de um ser humano. A inteligência artificial fraca se verifica quando o agente apresenta alguns traços de inteligência semelhantes aos de seres humanos, porém, em um campo ou tarefa específica. Esse é o caso do *chatbot*, que se limita a responder perguntas sobre um determinado assunto para o qual foi treinado.

Nas palavras de Jordan (2019):

Artificial Intelligence (AI) is the mantra of the current era. The phrase is intoned by technologists, academicians, journalists, and venture capitalists alike. As with many phrases that cross over from technical academic fields into general circulation,

there is significant misunderstanding accompanying use of the phrase. However, this is not the classical case of the public not understanding the scientists—here the scientists are often as befuddled as the public. The idea that our era is somehow seeing the emergence of an intelligence in silicon that rivals our own entertains all of us, enthralling us and frightening us in equal measure. And, unfortunately, it distracts us.¹

Às vezes, as pessoas tendem a compreender a expressão “inteligência artificial” como um sinônimo de “inteligência humana”. No entanto, ainda que alguns agentes sejam inteligentes e autônomos em certos aspectos, provavelmente permanecerão máquinas inconscientes ou sendo utilizados como dispositivos que apoiem os seres humanos em tarefas complexas e específicas. A realidade de um processamento digital é diferente daquela do processamento biológico, das qualidades e das habilidades cognitivas que se encontram nas criaturas, como nos seres humanos e em outros animais (Aïmeur, 2021).

Os algoritmos com inteligência artificial ainda não conseguem criar representações semânticas, não se engajam profundamente e criativamente, não têm emoções e não fazem inferências em condições de colocá-las como idênticas às de um ser humano. Ao mesmo tempo, os algoritmos podem ser desvirtuados para refletir fenômenos sociais complexos e danosos, que necessitam da sensibilidade humana para análise, numa contraposição à tomada de decisão produzida a partir da análise meramente objetiva.

Além disso, vícios como vieses e discriminações algorítmicas potencializam lesões a direitos e possuem a capacidade de naturalizar a violação de direitos e de garantias constitucionalmente asseguradas. Isso acontece porque tende a haver uma percepção de credibilidade absoluta no processamento computacional, principalmente em decorrência da associação de que

1 “A Inteligência Artificial (IA) é o mantra da era atual. A frase é entoada por tecnólogos, acadêmicos, jornalistas e capitalistas de risco. Tal como acontece com muitas frases que passam dos campos acadêmicos técnicos para a circulação geral, há um equívoco significativo acompanhando o uso da frase. No entanto, este não é o caso clássico de o público não entender os cientistas – aqui os cientistas são muitas vezes tão confusos quanto o público. A ideia de que nossa era está de alguma forma vendo o surgimento de uma inteligência em silício que rivaliza com a nossa nos diverte, encanta e assusta em igual medida. E, infelizmente, isso nos distrai.” (tradução nossa)

as decisões emitidas por máquinas seriam mais confiáveis por estarem desprovidas de subjetividades humanas.

A seguir, na Subseção 2.2, apresenta-se uma revisão da literatura da área. Por meio do estudo, percebe-se que mesmo com a dificuldade de se obter acesso à estrutura e à codificação com que são implementados os algoritmos de inteligência artificial, nas análises das decisões por eles produzidos, encontram-se reflexos das mesmas práticas violadoras de direitos que acontecem nas sociedades no plano fático.

2.2 A utilização da inteligência artificial na violação de direitos no processo penal

A revisão da literatura da área mostra que mesmo nas sociedades mais evoluídas tecnologicamente, nascedouro ou aperfeiçoadoras das mais inovadoras técnicas computacionais, a tomada de decisão orientada por mecanismos que se utilizam de estratégias de emulação do comportamento humano pode resultar na potencialização de políticas públicas e de políticas privadas discriminatórias. Além disso, pode incentivar a coisificação do indivíduo e a instrumentalização do ser humano, produzindo sociedades disfuncionais.

Um dos cenários de potencial aplicação de viés discriminatório pelos algoritmos que se utilizam de inteligência artificial é no campo da persecução penal. No trabalho de Yarovenko, Shapovalova e Ismagilov (2021), os autores destacam as possíveis consequências problemáticas do uso de ferramentas modernas de hardware e de software e dos métodos de fixação de penas e de reconhecimento facial pelo Estado.

Conforme o estudo destaca, o reconhecimento facial pode auxiliar as autoridades na localização de pessoas procuradas e na confirmação de autoria e de materialidade de crimes, porque permite objetivar e otimizar a produção da prova. Nesse sentido, a inteligência artificial, em conjunto com outras técnicas, permite analisar características, construir perfis e revelar relacionamentos entre diversas variáveis e atores de delitos.

No entanto, ao mesmo tempo em que se aumenta a capacidade preventiva e punitiva do Estado, abre-se a possibilidade de se violar o direito dos cidadãos sobre os seus dados pessoais armazenados – com ou sem o seu consentimento – nas bases estatais de segurança pública. Outra das várias possibilidades de lesão a direitos está na criação de perfis criminosos que concluem pelo potencial de lesividade de um indivíduo a partir de características físicas, emocionais e sociais, aumentando-se a atuação do Estado sobre núcleos já estigmatizados ou discriminados numa determinada sociedade.

Por esse motivo, é importante que o Estado defina regras por meio de legislação específica para estabelecer legitimidade ao reconhecimento facial na segurança pública. Considerando as amplas consequências sociais das ofensas, as oportunidades de abuso estatal na persecução penal e o desvirtuamento da ação, Shapiro (2019) discorre sobre o tema analisando que sistemas preditivos na persecução penal podem operar sobre as práticas sociotécnicas das autoridades e que essas podem concretizar uma atuação contraditória do Estado ao tentar promover a segurança produzindo danos.

A Comissão Europeia já se manifestou pelo desenvolvimento de uma inteligência artificial – com extensibilidade às demais técnicas computacionais modernas de manipulação de dados – dentro da legalidade, projetada e implementada de modo a refletir um processamento de dados ético e tecnicamente robusto, que respeite os valores democráticos, os direitos humanos e o Estado de Direito. A inteligência artificial aplicada no reconhecimento facial, cada vez mais utilizada em sistemas de monitoramento e de vigilância em massa, pode ser explorada de maneira a exacerbar a discriminação.

Conforme discorre Barrett (2020):

Attempts to infer personal attributes or emotions from someone's facial expression also invite privacy invasions and discrimination.⁸⁷ The long and ugly history of pseudoscientific attempts to connect physical appearance to mental and

moral aptitude will not be improved or corrected by incorporating those methods into algorithms. Systems that promise to “assess criminality” or assess a job applicant’s candidacy for the position will only reify existing inequality by providing a supposedly scientific justification for discrimination.⁸⁸ Studies have found that (...) these systems introduce an additional form of racial bias by misinterpreting the facial expressions of Black people,⁹⁰ generally providing them with more negative scores on average than people of other ethnicities.²

Cada vez mais a análise preditiva e a inteligência artificial são utilizadas na persecução penal, e essas técnicas têm impactos profundos na vida das pessoas envolvidas, como se verifica na aplicação de benefícios no cumprimento da pena pelo juízo da execução penal. Na direção dos trabalhos que defendem a aplicação dessas técnicas (Siegel, 2018; Corbett-Davies; Goel; González-Bailón, 2017), muito se divulga sobre a potencialidade de a análise automatizada de dados tornar a tomada de decisão no processo penal mais objetiva, consistente, rigorosa, neutra e sem a colocação de valoração pessoal na decisão.

No entanto, conforme refere Shapiro (2019), ao se investigar criticamente esses sistemas, verifica-se que essas são reivindicações de justiça que não se sustentam. A análise preditiva que se utiliza de dados já enviesados torna essa análise limitada e tendenciosa, fazendo com que o poder público volte a atenção da atuação policial para as comunidades que já são mais policiadas, enquanto causas subjacentes do crime – como a falta de investimento naquelas pessoas e a pobreza – são ignoradas.

Os Estados Unidos é um país que utiliza muito a inteligência artificial no sistema judiciário local – são mais de sessenta ferramentas de avaliação de risco utilizadas atualmente. Essas

2 “As tentativas de inferir atributos pessoais ou emoções a partir da expressão facial de alguém também convidam a invasões de privacidade e discriminação.⁸⁷ A longa e feia história de tentativas pseudocientíficas de conectar a aparência física à aptidão mental e moral não será melhorada ou corrigida pela incorporação daqueles métodos em algoritmos. Sistemas que prometem ‘avaliar a criminalidade’ ou avaliar a candidatura de alguém a uma vaga de trabalho só irão reafirmar a desigualdade existente ao fornecer uma justificativa supostamente científica para a discriminação.⁸⁸ Estudos descobriram que (...) esses sistemas introduzem uma forma de preconceito racial ao interpretar mal as expressões faciais de pessoas negras,⁹⁰ geralmente fornecendo-lhes pontuações mais negativas, em média, do que as pessoas de outras etnias.” (tradução nossa)

ferramentas alcançam desde a etapa policial, na identificação, na previsão e na intervenção nos locais com maior probabilidade de acontecerem crimes até a vigilância em tempo real. No trabalho de Angwin *et al.* (2016), os autores demonstram como um algoritmo usado para determinar a fiança era mais propenso a classificar acusados afro-americanos como pessoas de alto risco do que os acusados brancos.

Uma das realidades que já se verifica é a da discriminação geográfica, quando algumas áreas nas quais são aplicadas técnicas de inteligência artificial passam a ser mais monitoradas e controladas do que outras, o que resulta em mais prisões num determinado cenário local. No trabalho de Babuta e Oswald (2019), os autores mostram como a inteligência artificial pode fazer com que pessoas de certas comunidades sejam paradas e revistas mais frequentemente do que as demais. Em ambos os casos, as pessoas estão sendo privadas dos princípios de justiça, igualdade e equidade.

Nos últimos cinco anos, a imprensa também passou divulgar mais sobre como os algoritmos de aprendizado de máquina e a inteligência artificial podem ser impregnadas dos preconceitos dos que desenvolvem os seus algoritmos ou da sociedade no qual ele será aplicado. Nesse sentido, os textos de jornais de Crawford (2016); Lohr (2018); WachterBoettcher (2017).

Dedica-se uma atenção maior ao processo penal neste trabalho porque a restrição de liberdade é uma das mais severas penalidades de um sistema judicial. É exatamente na esfera penal que se tem visto o maior investimento estatal e privado no mundo, na criação de bases de dados com informações pessoais assim como no desenvolvimento de técnicas e de ferramentas de tomada de decisão que operem e concluam sobre tais dados, a fim de atingir um determinado objetivo.

A dificuldade para se enxergar e questionar a configuração explícita da discriminação algorítmica acontece porque tudo isso se desenvolve num contexto de baixa ou quase nenhuma transparência. Isso dificulta a demonstração de como o algorit-

mo pode estar estruturado para concluir – ou para direcionar o processo de aprendizagem a concluir – por respostas que carregam nelas mesmas a violação de direitos e de garantias. Assim, pode-se direcionar a tomada de decisão para uma determinada resposta que não corresponde ao correto, nem ao que implemente as disposições do ordenamento jurídico ou o que efetive os direitos humanos, mas ao desejado por um núcleo num certo contexto, como o político ou o social.

No trabalho de Rodríguez (2020), o autor explora algumas ferramentas desenvolvidas com inteligência artificial, que são aplicadas no sistema judiciário da Espanha. O pesquisador mostra como uma ferramenta, ao operar sobre parâmetros tendenciosos e subjetivos, tende a determinar que indivíduos de uma certa etnia apresentam um maior risco criminal do que indivíduos de outra. A presença de parâmetros racistas e desiguais torna o processo de decisão uma expressão de discriminação, e não uma ferramenta de concretização ou de promoção da justiça.

No momento em que um processo de análise algorítmica é instituído para operar sobre uma base de dados não condizente com a realidade, limitada ou incorreta, provavelmente, a conclusão a que se chegará será distorcida ou falsa. Isso possibilita a violação de direitos, incluindo a proteção dos dados pessoais, o direito à privacidade e o direito à não discriminação.

Uma das maiores preocupações do uso da inteligência artificial no processo de decisão judicial na esfera penal está na substituição da figura do juiz pelo julgamento dos indivíduos por meio de algoritmos e de máquinas sem consciência (Aïmeur, 2021). Isso porque uma sentença não deve ser construída substanciada na simples consideração de aspectos objetivos, como a adequação de um fato a uma tipificação pelo teor do evento praticado.

A inteligência artificial não tem a sensibilidade humana que permite o processo criativo de capturar a semântica no caso concreto, de avaliar e imaginar os fatos considerando a complexa rede de detalhes que se escondem nos elementos de sintaxe, motivo

pelo qual não consegue adaptar a forma de construir as diferentes decisões (Andrés, 2019) a partir das mudanças de um cenário para o outro quando há circunstâncias nos autos tão importantes quanto a objetividade da definição do enquadramento dos fatos. Por isso, tende a produzir decisões fossilizadas (Fenoll, 2018).

O Parlamento Europeu já se manifestou pela necessidade de que a inteligência artificial e as demais técnicas computacionais de extração, de análise e de manipulação de dados tenham de respeitar plenamente os direitos fundamentais consagrados na Carta dos Direitos Fundamentais da União Europeia, assim como a Legislação da União Europeia sobre a privacidade e proteção de dados, especialmente a Diretiva da União Europeia nº 2016/680.

O Parlamento Europeu também já destacou a necessidade de que os algoritmos sejam desenvolvidos aplicando-se na codificação e em todo o ciclo de vida dos mesmos vários princípios básicos. Dentre os quais destacam-se a aplicabilidade algorítmica e a transparência, a rastreabilidade, a implementação obrigatória de constantes avaliações de impacto do algoritmo sobre os direitos fundamentais antes da aplicação ou da implantação de qualquer sistema que se utilize de inteligência artificial e o desempenho de auditorias obrigatórias nesses sistemas.

Conforme expressa Escobedo, Meza e Trejo (2020), as relações intrínsecas que estão presentes nos dados refletem estruturas por vezes não explícitas numa sociedade, como a discriminação. Por conseguinte, uma ferramenta que aprenda sobre tais dados tenderá a apresentar resultados igualmente discriminatórios. A União Europeia tem reafirmado para a sociedade que o ponto de partida e a premissa ética para se desenvolver e se utilizar a inteligência artificial é a dignidade humana e os direitos fundamentais.

No trabalho de Gutierrez (2021), o autor destaca que uma das situações que torna difícil o combate da discriminação algorítmica é o fato de que as pessoas costumam perceber aquilo que é automatizado – como as decisões produzidas por algoritmos – como mais justas do que as que são produzidas pelos seres humanos. Isso também deriva do fato de que as empresas tentam

introduzir no imaginário do público uma certeza de que aquilo que desenvolveram é objetivo na consecução do que se essas se propõem a entregar.

Um estudo realizado pelo Pew Research Center (2018) indicou que 40% das pessoas nos EUA acreditavam que algoritmos podiam tomar decisões sem preconceito. No entanto, os dados brutos com os quais esses algoritmos trabalham são extraídos do mundo fático, dentro de um contexto, e podem estar enviesados mesmo que originalmente sejam neutros.

Conforme explica O'Neil (2017, p. 10):

Do you see the paradox? An algorithm processes a slew of statistics and comes up with a probability that a certain person might be a bad hire, a risky borrower, a terrorist, or a miserable teacher. That probability is distilled into a score, which can turn someone's life upside down.³

Na mesma direção, escreve Eubanks (2018, p.12) sobre como pessoas de alguns grupos deixados à margem pelo Estado são reduzidas e discriminadas pelas características conhecidas, ou construídas, a partir das informações descobertas, levantadas e/ou tratadas pelos algoritmos. No texto da escritora:

Automated decision-making shatters the social safety net, criminalizes the poor, intensifies discrimination, and compromises our deepest national values. It reframes shared social decisions about who we are and who we want to be (...) And while the most sweeping digital decision-making tools are tested (...) where there are few expectations of political accountability and transparency, systems first designed for the poor will eventually be used on everyone. America's poor and working-class people have long been subject to invasive surveillance, midnight raids, and punitive public policy that increase the stigma and hardship of poverty.⁴

3 "Você vê o paradoxo? Um algoritmo processa uma série de estatísticas e apresenta uma probabilidade de que uma determinada pessoa possa ser uma má contratação, um tomador de empréstimo arriscado, um terrorista ou um professor miserável. Essa probabilidade é destilada em uma pontuação, que pode virar a vida de alguém de cabeça para baixo". (tradução nossa)

4 "A tomada de decisão automatizada destrói a rede de segurança social, criminaliza os pobres, intensifica a discriminação e compromete nossos valores nacionais mais profundos. Ele reformula as decisões sociais compartilhadas sobre quem somos e quem queremos ser (...) E enquanto as ferramentas digitais de tomada

Conforme demonstrado, a inteligência artificial permite a reprodução de cenários de discriminação semelhantes aos que se encontram nas sociedades atuais. Os algoritmos de inteligência artificial são comumente utilizados em conjunto com outras técnicas que permitem a análise, a extração e a manipulação de imensas quantidades de dados pessoais em segundos, viabilizando a concretização de cenários de discriminação individual e coletiva, com a possibilidade, ainda, da indeterminação da responsabilização. A seguir, são apresentadas as considerações finais do trabalho, fortalecendo-se a necessidade de se definir e de se efetivar princípios universais no desenvolvimento dos algoritmos e na utilização das técnicas computacionais a fim de se evitar lesões a direitos.

3 CONCLUSÃO

A tecnologia tem como característica, decorrente do próprio ciclo de desenvolvimento, a projeção de cenários futuros, antecipando-se ao evento não desejado, a fim de se ajustar a codificação para evitá-lo. Para tanto, existe a fase de testes. Em direção distinta, no Direito, se espera o fato acontecer para então se verificar a incidência da norma. Isso faz com que regras sejam instituídas tardiamente, após o estado fático de lesão de direitos já estar configurado.

Quando se trata de matéria computacional – diante da alta capacidade de armazenamento de dados, da velocidade de processamento, da ampla possibilidade de compartilhamento e da descentralização do controle da informação –, torna-se quase impossível restabelecer às vítimas o estado anterior à lesão aos seus direitos. Pode se ter uma pretensa ideia de controle sobre a propagação do dano, mas não se o tem de fato.

Desde que a inteligência artificial começou a ser pensada, algumas décadas atrás, houve um desenvolvimento técnico-cien-

de decisão mais abrangentes são testadas (...) onde há poucas expectativas de responsabilidade política e transparência, os sistemas projetados inicialmente para os pobres acabarão sendo usados para todos. As pessoas pobres e da classe trabalhadora dos Estados Unidos há muito estão sujeitas à vigilância invasiva, incursões policiais de madrugada e políticas públicas punitivas que aumentam o estigma e as dificuldades da pobreza.” (tradução nossa)

tífico muito acelerado. Isso permitiu que a sociedade se desenvolvesse num movimento que, mesmo sem alguma regulamentação mais centralizada sobre aquelas tecnologias emergentes, fosse de algum modo limitado (quanto a alguns dos potenciais riscos de lesão a direitos) pelas próprias características da tecnologia existente à época e pelo cenário social.

No entanto, com o avanço do tempo e devido à forma como as sociedades foram se organizando, com as fronteiras econômicas e políticas sendo cada vez menores e mais convergentes entre si, as técnicas computacionais deixaram de ser objetivadas para criar ou animar coisas e passaram a ser aplicadas de maneira mais direta e invasiva sobre o ser humano. Dessa forma, os Estados passaram a experimentar o uso das técnicas computacionais como uma maneira de se exercer controle sobre a população.

Esse mesmo controle passou a ser visto pela iniciativa privada como uma forma de se fazer dinheiro. Com isso, enquanto a capacidade de armazenamento de dados aumentava, a internet ia alcançando um maior número de Estados e de estratos sociais, a velocidade de processamento seguia uma curva ascendente e as ferramentas computacionais iam sendo desenvolvidas sem a obrigação de satisfação de princípios ou de regulamentações. Então percebeu-se que o maior ativo econômico que havia era o próprio ser humano porque ele era uma fonte de dados.

Numa estrada sem obstáculos, a caixa preta da codificação permitia o desenvolvimento algorítmico marcado pela liberdade da violação de direitos. Assim, sem que forças direcionassem os valores locais ou universais que o processamento algorítmico deveria implementar – e se comprometer a efetivar – no ciclo de vida e no próprio código, abriu-se a possibilidade da objetificação do homem.

O ser humano como objeto, como coisa, o ser humano como instrumento, mas dentro da legalidade. Essa visão pautou a pesquisa científica de várias áreas por muitas décadas. Foi assim que grandes atrocidades foram praticadas na humanidade.

Esse é um ponto interessante: quando há lacunas na norma, ou vácuos de eficácia normativa, há a possibilidade de se lesionar gravemente direitos e garantias fundamentais. De maneira afrontosa à dignidade humana, pode se preencher os espaços e se subverter a proteção dos direitos previstos nos dispositivos constitucionais.

Isso é perigoso principalmente quando uma sociedade enfrenta crises econômicas, porque a pobreza amplifica o estado de necessidade. Onde falta a possibilidade do exercício da vida com dignidade, abre-se o espaço para a internalização de ideias que atingem o cerne da democracia e que desmontam a necessidade de se efetivar a dignidade humana.

Quanto a isso, diferentemente do que alguns pretendem fazer crer, da existência de normas e da instituição de fiscalização sobre o cumprimento dessas não deriva a implicação de retrocesso de desenvolvimento, tampouco de elevação de custos. Pelo contrário. Verifica-se que o progresso se efetiva no real sentido da palavra porque o índice de desenvolvimento humano tende a aumentar ao mesmo tempo em que se impede o retrocesso social e histórico das práticas violadoras de direitos humanos.

O volume de dados humanos a que se tem acesso hoje por meio das bases de dados públicas e das bases de dados privadas é da ordem de bilhões de registros. O avanço das técnicas computacionais permite a extração, o rápido cruzamento dessas informações, a inferência e a predição a tal ponto que o indivíduo se torna um ser consumido por modelos estatísticos e econômicos que valoram os dados pessoais como um dos principais ativos. Muito mais do que isso: o ser humano em si é transformado numa unidade de dados consumível.

Sobre dados em que se aplicam procedimentos e funções objetivas para se concluir quem é o indivíduo no mundo, tem-se o desaparecimento da perspectiva produzida pela sensibilidade humana. A mesma sensibilidade com que se construiu a ideia universal dos direitos humanos.

Por meio da sensibilidade aplicada à racionalidade, e não da objetificação, foi que se construiu o complexo processo de socialização humana com ênfase em direitos universais. Esses direitos nos permitiram evoluir ao ponto de se conviver em sociedade com um mínimo padrão de igualdade – até mesmo de segurança na escala de valorização da vida de cada ser humano – a ser observado por todos.

Nesse sentido, quando o ser humano é reduzido a um conjunto de dados disponíveis para manipulação – e dados são elementos manipuláveis –, o ser humano está sendo igualmente manipulado. E se não se depende mais dele como pessoa natural, mas de um conjunto de dados para se definir quem ele é diante das regras de quem o avalia, ainda que seja o algoritmo o avaliador, então esse indivíduo tem seu espaço de construção de identidade pessoal e social violado.

Ele não é mais quem ele sabe ser nem é entendido ou contado pela história que constrói. Ele passa a ser o que o algoritmo diz que ele é. O indivíduo perde não só a capacidade de se autodeterminar em muitas de suas decisões, como passa a ter um significado no mundo construído não mais a partir de si mesmo, mas a partir das conclusões dos outros, das conclusões a que chegaram algorítmicamente sobre ele.

Isso pode ter consequências intra e extra indivíduo avassaladoras, principalmente quando, deliberadamente, apoia-se a decisão na resposta algorítmica. Nesses casos, a tomada de decisão acontece sem qualquer análise humana permeada de sensibilidade, e os algoritmos não apresentam a habilidade e a flexibilidade necessária para o entendimento de contextos.

Essa situação é igualmente preocupante quando acontecem as parcerias entre Estado e empresas privadas, porque podem representar a violação de direitos legalizada em prol de um pretensão benefício comum. O que se verifica é que, nesses casos, muitas vezes não se alcança os ganhos que a sociedade teria com a atuação ou com o investimento do próprio Estado se ele agisse pautado nos princípios a que se vincula o exercício da administração pública.

Ao mesmo tempo, quanto mais projeção e controle sobre os perfis dos cidadãos um Estado possui, mais ele se direciona a alcançar um estado de potencial violação de direitos civis quando as bases democráticas começam a ser ruídas ou quando são efetivamente destruídas. Nesse sentido, nem toda a ausência de norma com real eficácia na tutela de direitos significa uma dificuldade do poder público em tratar um tema. Pode ser uma estratégia para que as violações de direito aconteçam sob o olhar complacente do Estado, a fim de beneficiar os interesses particulares de alguns.

Concluiu-se com a revisão da literatura da área que mesmo nas sociedades mais evoluídas tecnologicamente, verifica-se que a inteligência artificial pode ser utilizada para (re)produzir graves violações a direitos, potencializando a estigmatização, a coisificação e a instrumentalização do ser humano. Conforme demonstrado neste artigo, as correlações e as causalidades estabelecidas entre dados objetivos, subjetivos, sensíveis ou inferidos oportunizam graves lesões aos direitos.

Há de se pensar no desenvolvimento de algoritmos que promovam a inovação, mas que ao mesmo tempo respeitem os valores e os direitos fundamentais, assim como atendam aos princípios éticos. Ainda, com a possibilidade da determinação da responsabilidade pelo que se conclui e efetiva e nos quais haja transparência sobre as regras de processamento.

Não raro, as conclusões algorítmicas – seja pela inconsistência de dados, pelos dados enviesados ou pelo direcionamento tendencioso da codificação, entre outras situações – fundamentam muitas das atuações estatais violadoras de direitos (perpetradas diretamente pelo Estado ou pelos agentes que a ele respondem). Conforme demonstrado, alguns dos algoritmos usados na persecução penal norte-americana se utilizam de dados das bases imensas dos sistemas de dados dos diferentes órgãos do país, e dos processos judiciais, para apoiar ou, ainda, substituir a tomada de decisão humana na persecução penal.

Algumas daquelas ferramentas permitem concluir pela concessão de liberdade ao indivíduo, analisando elementos como

as dificuldades financeiras, os problemas vocacionais ou educacionais, a criminalidade familiar, o meio ambiente, a instabilidade residencial e a personalidade. Essas situações aumentam o preconceito, a estereotipagem e a vulnerabilidade das pessoas.

As possíveis violações de direitos podem levar os indivíduos a desistir de buscar a efetivação dos seus direitos pelos riscos de uma exposição social prejudicial e dos demais resultados danosos decorrentes de um dia terem buscado a tutela jurisdicional. Assim, diminui-se a força normativa da Constituição e desacreditam-se as bases do Estado Democrático de Direito, ao mesmo tempo em que o Estado olha de maneira complacente para as violações de direitos e de garantias fundamentais.

A potencialidade da inteligência artificial e das demais técnicas computacionais modernas de análise, de extração e de manipulação de dados cria um cenário propício para a (re)produção ou para a (re)criação das subjetividades excludentes existentes nas diferentes sociedades, que são transportadas do mundo fático para um contexto algorítmico. Assim, quanto mais disfuncional é uma sociedade na proteção dos direitos das pessoas, maior a probabilidade de que os algoritmos sejam desenvolvidos, aprendam ou expressem resultados compatíveis com a discriminação praticada na realidade local ou que sejam compatíveis com determinada diferenciação que alguns queiram promover.

Essas situações podem aumentar o preconceito, a estereotipagem e a vulnerabilidade das pessoas na persecução penal. Não se está dizendo que a inteligência artificial ou que qualquer outra técnica é por si mesma um problema, mas que há de se determinar que a criação e a exploração de seu uso aconteça dentro de um espaço de respeito aos direitos humanos e de efetivação dos direitos dos indivíduos expressos no ordenamento jurídico pátrio. Para tanto, entende-se que não bastam apenas regramentos (leis), é preciso pensar em algo maior, como em princípios universais que pautem todo o ciclo de vida das ferramentas computacionais, mas isso é assunto para outro artigo. ❖

REFERÊNCIAS

AÏMEUR, Esmá. Human versus artificial intelligence. *In: FRONTIERS IN ARTIFICIAL INTELLIGENCE*, Soesterberg, 25 mar. 2021. Disponível em: <https://www.frontiersin.org/articles/10.3389/frai.2021.622364/full>. Acesso em: 25 abr. 2022.

ANDRÉS, Núria. La verdad y la ficción de la inteligencia artificial en el proceso penal. *In: FUENTES, Jesús; HOYO, Gregorio (coords.). La justicia digital en España y la Unión Europea*. Barcelona: Atelier, 2019. p. 31-39.

ANGWIN, Julia; JEFF, Larson; MATTU, Surya; KIRCHNER, Lauren. Machine Bias: there's software used across the country to predict future criminals. And it's biased against blacks. *In: PROPUBLICA*. Chicago, 23 mai. 2016. Disponível em: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. Acesso em: 25 abr. 2022.

BABUTA, Alexander; OSWALD, Marion. Data analytics and algorithmic bias in policing. *In: ROYAL UNITED SERVICES INSTITUTE FOR DEFENCE AND SECURITY STUDIES*. Londres, 15 jul. 2019. Disponível em: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/831750/RUSI_Report_-_Algorithms_and_Bias_in_Policing.pdf. Acesso em: 25 abr. 2022.

BARRETT, Lindsey. Ban facial recognition technologies for children and for everyone else. Boston University **Journal of Science and Technology Law**. Boston, v. 26, n. 2, p. 223-285, jul. 2020. Disponível em: <https://www.bu.edu/jostl/files/2020/08/1-Barrett.pdf>. Acesso em: 25 abr. 2022

BARTNECK, Christoph; LÜTGE, Christoph; WAGNER, Alan; WELSH, Sean. What is AI? *In: BARTNECK, Christoph; LÜTGE, Christoph; WAGNER, Alan; WELSH, Sean (coords.). An Introduction to Ethics in Robotics and AI*. Cham: Springer, 2020. p. 5-16. Disponível em: https://link.springer.com/chapter/10.1007/978-3-030-51110-4_2. Acesso em: 25 abr. 2022.

CORBETT-DAVIES, Sam; PIERSON, Emma; FELLER, Avi; GOEL, Sharad. A computer program used for bail and sentencing decisions was labeled biased against blacks. It's actually not that clear. **Washington Post**, Washington, 17 out. 2016. Disponível em: <https://www.washingtonpost.com/news/monkey-cage/wp/2016/10/17/can-an-algorithm-be-racist-our-analysis-is-more-cautious-than-propublicas/>. Acesso em: 25 abr. 2022.

CORBETT-DAVIES, Sam; GOEL, Sharad, GONZÁLEZ-BAILÓN, Sandra. Even imperfect algorithms can improve the criminal justice system. **New York Times**, Nova York, 20 dez. 2017. Disponível em: <https://www.nytimes.com/2017/12/20/upshot/algorithms-bail-criminal-justice-system.html#:~:text=the%20main%20story-Even%20Imperfect%20Algorithms%20Can%20Improve%20the%20Criminal%20Justice%20System,biased%20nature%20of%20human%20decisions>. Acesso em: 25 abr. 2022.

CRAWFORD, Kate. Artificial intelligence's white guy problem. **New York Times**, Nova York, 25 jun. 2016. Disponível em: <https://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html>. Acesso em: 25 abr. 2022.

ESCOBEDO, Fernanda; MEZA, Iván; TREJO, Sofía. Hacia los comités de ética en inteligencia artificial. In: ARXIV - CORNELL UNIVERSITY. Ithaca, 11 fev. 2020. Disponível em: <https://arxiv.org/ftp/arxiv/papers/2002/2002.05673.pdf>. Acesso em: 25 abr. 2022.

EUBANKS, Virginia. **Automating inequality: how high-tech tools profile, police, and punish the poor**. Nova York: St. Martin's Press, 2018. p.12.

EUROPA. Parlamento Europeu. **Propuesta del Parlamento Europeo sobre la inteligencia artificial en el derecho penal y su utilización por las autoridades policiales y judiciales en asuntos penales**. Bruxelas, BE: Parlamento Europeu, 2020. Disponível em: https://www.europarl.europa.eu/doceo/document/A-9-2021-0232_ES.html. Acesso em: 25 abr. 2022.

EUROPA. Comissão Europeia. **Livro Branco sobre a inteligência artificial - uma abordagem europeia virada para a excelência e a confiança**. Bruxelas, BE: Comissão Europeia, 2020. Disponível em: https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_pt.pdf. Acesso em: 25 abr. 2022.

FENOLL, Jordi. **Inteligencia artificial y proceso judicial**. Madri: Marcial Pons Ediciones Jurídicas y Sociales, 2018.

GUTIERREZ, Miren. Algorithmic gender bias and audiovisual data: a research agenda. **International Journal of Communication**, [s. l.], v. 15, n. 1, p. 439-461, jan. 2021. Disponível em: <https://ijoc.org/index.php/ijoc/article/viewFile/14906/3333>. Acesso em: 25 abr. 2022.

JORDAN, Michael. Artificial intelligence - the revolution hasn't happened yet. In: HARVARD DATA SCIENCE REVIEW – MIT PRESS, Cambridge, 01 jul. 2019. Disponível em: <https://hdr.mitpress.mit.edu/pub/wot7mkc1/release/9>. Acesso em: 25 abr. 2022

KAPLAN, Andreas; HAENLEIN, Michael. A brief history of artificial intelligence: on the past, present, and future of artificial intelligence, **California Management Review**, Califórnia, v. 61, n. 4, p. 5-14, ago. 2019. Disponível em: <https://journals.sagepub.com/doi/10.1177/0008125619864925>. Acesso em: 25 abr. 2022.

LOHR, Steve. Facial recognition is accurate, If you're a white guy. **New York Times**, Nova York, 09 fev. 2018. Disponível em: <https://www.nytimes.com/2018/02/09/technology/facial-recognition-race-artificial-intelligence.html>. Acesso em: 25 abr. 2022.

MÁNTARAS, Ramon. Towards artificial intelligence: advances, challenges, and risks. **Metode Science Studies Journal**, [s. l.], v. 1, n. 9, p. 119-125, mar. 2019. Disponível em: <https://ojs.uv.es/index.php/Metode/article/view/11145>. Acesso em: 25 abr. 2022.

O'NEIL, Cathy. **Weapons of math destruction**: how big data increases inequality and threatens democracy. Nova York: Broadway Books, 2017. p. 10.

POOLE, David; MACKWORTH, Alan. **Artificial intelligence foundations of computational agents**. Cambridge: Cambridge University Press, 2017.

RODRÍGUEZ, Ana. The impact of artificial intelligency on penal procedure. **Anuario de la Facultad de Derecho de la Universidad de Extremadura**, Cáceres, v. 1, n. 36, p. 695-728, dez. 2020. Disponível em: <https://publicaciones.unex.es/index.php/AFD/article/view/489>. Acesso em: 25 abr. 2022.

RUSSELL, Stuart; NORVIG, Peter. **Artificial intelligence**: a modern approach. New Jersey: Pearson. 2020.

SCHWAB, Klaus. **A quarta revolução industrial**. São Paulo: Edipro, 2016.

SHAPIRO, Aaron. Predictive policing for reform? Indeterminacy and intervention in big data policing. **Surveillance and Society**, [s. l.], v. 17, n. 3-4, p. 456-472, set. 2019. Disponível em: <https://doi.org/10.24908/ss.v17i3/4.10410>. Acesso em: 25 abr. 2022.

SIEGEL, Eric. How to fight bias with predictive policing. *In*: **SCIENTIFIC AMERICAN**. Nova York, 19 fev. 2018. Disponível em: <https://blogs.scientificamerican.com/voices/how-to-fight-bias-with-predictive-policing>. Acesso em: 25 abr. 2022.

SMITH, Aaron. Attitudes toward algorithmic decision making. *In*: **PEW RESEARCH CENTER**. Washington, 16 nov. 2018. Disponível em: <https://www.pewresearch.org/internet/2018/11/16/attitudes-toward-algorithmic-decision-making/>. Acesso em: 25 abr. 2022.

STRAUß, Stefan. From big data to deep learning: a leap towards strong AI or 'intelligentia obscura'?. **Big Data and Cognitive Computing**, [s. l.], v. 2, n. 3, p. 1-19, jul. 2018. Disponível em: <https://www.mdpi.com/2504-2289/2/3/16/pdf>. Acesso em: 25 abr. 2022.

WACHTER-BOETTCHER, Sara. How Silicon Valley's blind spots and biases are ruining tech for the rest of us. **Washington Post**, Washington, 13 dez. 2017. Disponível em: <https://www.washingtonpost.com/news/posteverything/wp/2017/12/13/how-silicon-valleys-blind-spots-and-biases-are-ruining-tech-for-the-rest-of-us/>. Acesso em: 25 abr. 2022.

YAROVENKO, Vasily; SHAPOVALOVA, Galina; ISMAGILOV, Rinat. Some problems of using the facial recognition system in law enforcement activities. **Правовое государство теория и практика**, [s. l.], v. 17, n. 1, p. 189-200, mar. 2021. Disponível em: <https://pravgos.ru/index.php/journal/article/view/160>. Acesso em: 25 abr. 2022.