INTELIGÊNCIA ARTIFICIAL NA SEGURANÇA PÚBLICA: APLICAÇÕES E CONSIDERAÇÕES ÉTICAS

Jordana Martins Perussi[1], Ana Beatriz da Silva[2], Paulo Cesar Correa Borges[3] e Patrícia Borba Marchetto[4].

RESUMO

O presente artigo visa tratar sobre a aplicação da inteligência artificial no âmbito da segurança pública, além das possíveis repercussões éticas ao redor desse tema. Para isso, utilizando-se do método dedutivo, foi realizada a pesquisa, leitura e revisão bibliográfica específica sobre o tema, com o intuito de conceituar a inteligência artificial, delimitar sua abrangência e aplicação no âmbito da segurança pública. Concluiu-se que o uso dessa tecnologia, e de suas subáreas, ao se tratar de segurança pública, tem sido feito principalmente em relação ao reconhecimento facial e ao policiamento preditivo. Ainda, realizou-se um estudo de caso da cidade de Chicago, nos Estados Unidos. Por fim, ressalta-se que o uso desses sistemas deve ser feito de forma cautelosa, vez que estes podem refletir uma falsa percepção de objetividade, quando, na realidade, estão intrinsecamente ligados com a forma que foram programados.

Palavras-chave: Inteligência Artificial; Segurança Pública; Ética.

ABSTRACT

This article aims to address the application of artificial intelligence in the field of public security, as well as the possible ethical repercussions surrounding this issue. To this end, using the deductive method, research, reading and a specific literature review on the subject were carried out, with the purpose of conceptualizing artificial intelligence, delimiting its scope and application within the scope of public security. It was concluded that the use of this technology, and its sub-areas, when it comes to public security, has been made mainly in relation to facial recognition and predictive policing. A case study of the city of Chicago in the United States was also made. Finally, it should be emphasized that the use of these systems should be done with caution, as they can reflect a false perception of objectivity, when in reality they are intrinsically linked to the way they have been programmed.

Keywords: Artificial Intelligence; Public Security; Ethics.

[1] Mestranda em Direito pela UNESP, jordana.martins@unesp.br;
[2] Graduanda em Direito pela UNESP, ana-beatriz.silva@unesp.br;
[3] Orientador, docente da UNESP, paulo.cc.borges@unesp.br
[4] Orientadora, docente pela UNESP, patricia.marchetto@unesp.br

1. INTRODUÇÃO

É perceptível que, nos últimos anos, uma nova era tecnológica se estabeleceu afetando os indivíduos em todos os âmbitos de suas vidas. Tais inovações emergem acompanhadas de dilemas éticos que englobam tanto sua aplicação, quanto uso, razão pela qual debates quanto os desdobramentos e reflexos que essa nova tecnologia pode gerar merecem destaque. Dessa forma, tecnologias como a inteligência artificial (IA), e suas subáreas, se enquadram nesse contexto, ao passo em que se propagam e se expandem.

Assim, o presente trabalho possui como objetivo central analisar os impactos e possíveis desdobramentos, principalmente éticos, da IA no âmbito da segurança pública. Nesse sentido, com o intuito de realizar um estudo mais aprofundado, é necessário identificar os conceitos de inteligência artificial, machine learning, deep learning e big data. Motivo pelo qual, se realizou uma revisão bibliográfica em relação a textos específicos quanto ao tema.

Ademais, se tratando da aplicação da IA e seus subgrupos quando se trata de segurança pública, observou-se que esta é manejada em relação ao uso do reconhecimento facial e do policiamento preditivo. Ainda, analisando de forma crítica, nota-se que sistemas a base de IA são capazes de transmitir uma falsa percepção de imparcialidade em seus resultados. Entretanto, não restam dúvidas de que as conclusões trazidas por esses mecanismos, são reflexos de sua programação, bem como dos danos com os quais foram “alimentados”.

Visando exemplificar o exposto, será feito o estudo de caso referente a cidade de Chicago, nos Estados Unidos, a qual fez a implementação de uma Strategic Subject List (SSL). Essa ferramenta possuía como intuito incorporar dados, classificar e identificar indivíduos que poderiam ser tornar vítimas ou possíveis perpetradores de tiroteios. Contudo, como será demonstrado, a aplicação de mecanismos como este deve ser acompanhada de discussões sobre os aspectos éticos, principalmente voltados ao tipo de dados que são coletados e como é realizada essa captação.

Por fim, denota-se que, apesar do uso da IA no âmbito da segurança pública demonstrar ser algo promissor, a aplicação deve ser feita com cautela, levando em consideração os possíveis prejuízos que podem gerar. Portanto, a incorporação de ideais de transparência e explicabilidade em relação a esses mecanismos, devem ser realizadas desde a programação até o uso.

2. A INTELIGÊNCIA ARTIFICIAL NO CAMPO DA SEGURANÇA PÚBLICA

2.1. Inteligência Artificial: Noções Gerais

Os avanços tecnológicos dos últimos anos trouxeram consigo novos dilemas éticos e principiológicos, os quais estão sendo objeto de grande debate no mundo acadêmico e na sociedade civil. Mecanismos como a inteligência artificial (IA) ganharam foco, ao passo que estão cada vez mais presentes no cotidiano das pessoas. Nesse sentido, com o intuito de melhor compreender as proporções que os sistemas de inteligência artificial estão tomando, é necessário antes estudar a origem e o conceito desse modelo.

Apesar de ser algo que ganhou popularidade nos últimos, o termo "inteligência artificial" foi estabelecido pela primeira vez em 1956, em um evento realizado em Dartmouth College, e o qual recebeu o nome de "Dartmouth Summer Research Project on Artificial Intelligence". A conclusão obtida pelos cientistas nesse encontro foi de que quaisquer tarefas descritas de forma precisa, poderiam ser reproduzidas por computadores programados. Mesmo com este ânimo inicial, a IA passou por um período mormente, em que muitos projetos não se concretizaram. (Kaufman, 2019).

Entretanto, recentemente o cenário mudou, principalmente, em decorrência do aumento do uso da internet. Pessoas ao redor do mundo adquiriram o hábito de passar muito tempo online, ainda mais no contexto da pandemia de COVID-19, emergindo até mesmo como um meio alternativo à presencialidade. Assim, quantidades massivas de dados começaram a ser gerados, o que permitiu que mecanismos, como a inteligência artificial, fossem utilizados com o intuito de sistematizar essas informações.

Sob esse prisma, o conceito de inteligência artificial passou a ser debatido, principalmente, no âmbito internacional e acadêmico. Contudo, é importante ressaltar que não há um consenso sobre qual seria a definição mais adequada para descrever esse mecanismo complexo. Por sua vez, a UNESCO em sua "Recomendação sobre a Ética da Inteligência Artificial", destaca até mesmo certa impossibilidade de se definir IA em razão da constante mudança pelo qual esses sistemas passam (2022).

Dessa forma:

> Não existe consenso na literatura científica em relação ao conceito de IA, mas suas definições costumam ser categorizadas empiricamente e teoricamente. O que significa dizer que, na ótica da categoria empírica, existe a perspectiva de os sistemas pensarem como seres humanos e aprenderem a partir da experiência.[...] Enquanto isso, no enfoque teórico

> esperam-se ações lógicas, que incluem capacidade de dedução e inferência sobre novas relações (Alves, 2020). O reconhecimento de padrões e aplicação de regras lógicas são elementos relevantes no modo de funcionamento das técnicas de IA. (Bottino; Fernanda 2023)

Independente de qual entendimento adotado sobre a categoria de funcionamento da IA, os algoritmos são elementos essenciais para o funcionamento desses sistemas operacionais. Assim, podem ser entendidos, de forma simples, como conjuntos de informações, claras, capazes de serem lidas e identificadas por um computador. Ou ainda, são um "conjunto de instruções matemáticas, uma sequência de tarefas para alcançar um resultado esperado em um tempo limitado". (Kaufman, 2019)

Sendo assim, observa-se que os algoritmos desempenham a função de organizar quantidades massivas de dados, além de conduzirem à descoberta de possíveis resoluções de problemas. Em decorrência disso, nota-se a maneira como as "aplicações de IA apenas são possíveis para a solução de problemas da área da ciência da computação a partir da disponibilidade abundante de dois insumos fundamentais: bancos de dados massivos e grande poder computacional." (Bottino;Fernanda, 2023)

Por fim, mesmo que a IA não possua uma definição permanente e precisa, ainda é possível extrair características que permitam reconhecê-la. Diante disso, ela pode ser compreendida como um complexo que mistura sistemas computacionais e de processamento de informações, integrando algoritmos capazes de realizar tarefas de caráter cognitivo (Russell; Norvig, 2022, apud Andrade, et. al. 2022).

2. 2. Machine Learning, Deep Learning e Big Data

Passando a analisar os desdobramentos decorrentes dos sistemas de inteligência artificial, o *machine learning* (ML) emerge como um modelo adotado frequentemente, sendo uma subárea da IA. O termo foi criado em 1959, pelo cientista Arthur Lee Samuel que partiu da ideia de que seria possível que computadores aprendessem sem serem programados. Dessa maneira, esse formato explora a construção de algoritmos que, programados, seguem instruções e são capazes de realizar previsões ou tomar decisões, a partir de dados (amostras) (Kaufman, 2019).

Na prática, e de forma simplificada, o ML ocorre quando"um computador observa alguns dados, monta um modelo baseado nos dados e usa o modelo como uma hipótese sobre o mundo é um software que pode resolver problemas" (Russell; Norvig, 2022). Esse sistema é utilizado em áreas da computação em que realizar a programação de algoritmos é

extremamente difícil ou até mesmo inviável. Ainda, possuem aplicabilidades em realizar grandes quantidades de previsões, o que não é possível para o ser humano antever. São por essas e outras questões que esses modelos passaram a ser mais utilizados, principalmente no cenário atual, marcado pela existência de volumes massivos de dados, que não são possíveis de serem processados a partir da programação computacional tradicional.

Tratando ainda do ML, Russel e Norving (2022), apontam a possibilidade de que esse modelo seja classificado em mais dois subgrupos: os de aprendizado supervisionado e o não supervisionado. No primeiro, o agente aprende com base na observação de dados de entrada e saída, recebendo e aplicando uma função capaz de realizar o mapeamento entre elas, visando, posteriormente, identificar padrões de saídas corretas. Por sua vez, o segundo recebe apenas os dados de entrada, sem receber uma função indicativa das saídas, ao passo em que deve analisar e identificar novos padrões.

O *deep learning*, por sua vez, é tido como uma espécie de *machine learning*, podendo ser definido como um "modelo estatístico de previsão de cenários futuros e a probabilidade de eles se realizarem e quando; a denominação provém da profundidade das camadas que formam a arquitetura das redes neurais." (Kaufman, 2022). Assim, esse modelo é capaz de lidar com dados de diversas dimensões e complexidade, como, por exemplo, o reconhecimento e processamento de diversos pixels em uma imagem.

Ainda, esse modelo utiliza como método as chamadas redes neurais artificiais (RNA), as quais têm como base a "arquitetura dos neurônios humanos e destinadas a reproduzir aprendizado por meio do desenvolvimento de sistemas que aprendem com exemplos de treinamento". (Bottino; Fernanda, 2023). Um exemplo de uso desses sistemas ocorre no campo da segurança pública, em que as RNA são aplicadas no policiamento preditivo, o qual faz uso do processamento de informações realizado pelos algoritmos, a partir de bancos de dados e com o intuito de realizar previsões capazes de auxiliar na efetividade da segurança pública. (Bottino; Fernanda, 2023)

Convém ressaltar por fim o conceito de Big Data, tendo em vista que este é categorizado como uma subárea que compõe as aplicações da inteligência artificial, conforme a Association for the Advancement of Artificial Intelligence (AAAI), referência na área. Não existe um consenso acerca do significado da expressão Big Data, a qual será melhor abordada mais adiante, todavia, pode ser pensada como uma quantidade massiva de dados armazenados, de forma estruturada ou não (Kaufman, 2019). Em razão disso, são necessárias ferramentas variadas capazes de decifrar essa quantidade de dados, ao mesmo tempo em que, de certa forma, permitem a visualização de novos padrões, capazes de melhor compreender a

sociedade. Possui relevância na medida em que o Big Data, como uma subárea da aplicação da inteligência artificial, é amplamente utilizado no âmbito da segurança pública.

2.3. O emprego da inteligência artificial na segurança pública

É notável o avanço do emprego de tecnologias no âmbito da segurança pública com o intuito de aprimorar investigações policiais e táticas de policiamento, buscando diminuir índices de criminalidade.

De acordo com Christopher Rigano (2019), o Instituto Nacional de Justiça dos Estados Unidos está focado no desenvolvimento de pesquisas de inteligência artificial voltadas à segurança pública, especialmente no que tange à análise de vídeos e imagens; análise de DNA; detecção de tiros e previsão de crimes. E, de maneira geral, destaca a importância do potencial da inteligência artificial na análise de grandes volumes informacionais.

Destaca o trabalho da Carnegie Mellon University com pesquisas voltadas à capacidade da máquina identificar o rosto do indivíduo em diferentes ângulos ou utilizando acessórios que impeçam a visualização completa da face como bonés, máscaras ou capacetes e ainda, em situações em que há dificuldade de iluminação ou imagens de baixa resolução (Rigano, 2019).

Na mesma esteira, destaca o trabalho dos pesquisadores da Dartmouth College, voltados ao desenvolvimento de algoritmos que degradam imagens de alta qualidade, com o intuito de fazer com que consigam reconhecer indivíduos ou placas de carro mesmo em vídeos ou imagens de qualidade inferior (Rigano, 2019).

Adentrando a esfera do emprego de tecnologias na segurança pública brasileira, destaca-se a pesquisa realizada pela Fundação Getúlio Vargas, cujos resultados estão presentes no relatório “Segurança Pública na era do Big Data” (2023). Referido relatório destaca a incorporação do big data nas estratégias empregadas pelas autoridades no combate à criminalidade, como por exemplo no tratamento de dados e na persecução penal. Na mesma esteira, ressalta as informações contidas no relatório Chicago Tonight apresentando dados acerca da redução do índice de criminalidade em 13% na cidade de Chicago, que conta com a incorporação de tecnologias de big data (Bottino; Fernanda, 2023).

De suma importância a compreensão do conceito de big data. O termo surge no século XXI, primordialmente nos ramos da astronomia e genética, com o intuito de expressar a ideia da necessidade de ferramentas capazes de processar e analisar extensos bancos de dados. Não existe um consenso acerca do significado da expressão big data, pois esta é bastante ampla

(Bottino; Fernanda, 2023). O relatório da Fundação Getúlio Vargas, para os fins de seus estudos, define big data como:

> A análise de grandes quantidades de dados, realizada de maneira automatizada por algoritmos, com intuito de extrair resultados e benefícios. Isso vislumbrando-se ainda que big data é menos sobre dados e mais sobre a capacidade de pesquisa, agregação e referência cruzada de grandes conjuntos de dados.

Dessa maneira, no âmbito da segurança pública, esta expressão, associada à ideia de uma quantidade massiva de dados, bem como sua extração, interpretação e análise, está diretamente vinculada à ferramentas de reconhecimento facial e policiamento preditivos, baseados em dados previamente coletados, fichas criminais e prévias informações de suspeitos.

Preliminarmente, o relatório destaca uma dificuldade em reunir informações envolvendo a contratação e emprego de tecnologias voltadas à segurança pública disponíveis ao público. Tais dificuldades se devem em razão do caráter sigiloso envolvendo investigações criminais, bem como restrições de compartilhamento de informações internas das instituições (Bottino; Fernanda, 2023).

Durante as entrevistas conduzidas com as autoridades para os fins do trabalho desenvolvido pela Fundação Getúlio Vargas, destacaram-se algumas tecnologias mencionadas de forma mais recorrente por serem empregadas na segurança pública brasileira, sendo elas: câmeras corporais, drones, reconhecimento facial, reconhecimento óptico de caracteres (OCR) e o policiamento preditivo (Bottino; Fernanda, 2023). Dentre essas, o reconhecimento facial e o policiamento preditivo serão analisados mais a fundo.

2.3.1. Reconhecimento facial

No âmbito nacional, destaca-se o emprego das tecnologias de reconhecimento facial especialmente nos estados do Acre, Amazonas, Pará, Roraima, Bahia, Minas Gerais, Goiás, Espírito Santos e Paraná (Bottino; Fernanda, 2023).

Costa e Kremer (2022) elucidam que o reconhecimento facial é uma tecnologia cujo funcionamento está ligado ao conceito de big data. Nesse sentido, explicam os autores:

> O reconhecimento facial é uma tecnologia de identificação biométrica realizada a partir da coleta de dados faciais, que podem ser provenientes de fotografias ou segmentos de vídeos. Esses sistemas automatizados extraem representações matemáticas de traços específicos como, por exemplo, a distância entre os olhos ou o formato do nariz, produzindo o que é chamado de padrão facial. É justamente no processo de comparação desse padrão

> facial a outros padrões faciais contidos na base de dados prévia do sistema que a tecnologia identifica indivíduos desconhecidos, como no caso das câmeras de monitoramento nas ruas [...].

Destacam ainda a existência de casos de prisões que ocorreram no Brasil em razão da utilização da tecnologia de reconhecimento facial, tal como um homem, procurado por homicídio, que foi identificado e preso no Carnaval de 2019 em Salvador. E ainda, a utilização de 28 câmeras munidas com reconhecimento facial no Carnaval do Rio de Janeiro que auxiliaram na identificação de quatro criminosos (Costa; Kremer, 2022).

No estado do Acre é possível mencionar o programa “Rio Branco mais segura”. A iniciativa contava com a instalação de 430 câmeras espalhadas pela cidade, sendo que destas, 18 estariam equipadas com a tecnologia de reconhecimento facial. Primordialmente, eram visados espaços públicos com grande movimentação de indivíduos, tais como o centro da cidade, Regional da Seis de Agosto, Parque Chico Mendes, Rodoviária Internacional de Rio Branco e o Horto Florestal (Bottino; Fernanda, 2023).

De importante destaque o estado de Roraima, o qual, em razão de sua localização, possui estreita ligação com a guerra às drogas e violência transfronteiriça. Em razão disso, a cidade de Pacaraima contou com a instalação de vinte luminárias inteligentes, as quais contêm câmeras de vigilância e tecnologia de reconhecimento facial (Bottino; Fernanda, 2023).

O estado da Bahia foi pioneiro no emprego do sistema de reconhecimento facial, e à época do estudo realizado pela Fundação Getúlio Vargas, a tecnologia já havia auxiliado em 351 prisões. O estado totaliza 78 municípios com a tecnologia, sendo 1.200 câmeras apenas em Salvador e região metropolitana (Bottino; Fernanda, 2023).

2.3.2 Policiamento preditivo

Através do levantamento realizado pela Fundação Getúlio Vargas, apenas dois estados brasileiros foram apontados como usuários da tecnologia de policiamento preditivo. O policiamento preditivo utiliza modelagem computacional aplicada a dados criminais passados com o intuito de prever atividades criminosas futuras (Bottino; Fernanda, 2023).

No mesmo sentido, Crawford, Richardson e Schultz (2019) esclarecem que a expressão “policiamento preditivo” está ligada a ideia de um sistema que executa a análise de dados com o intuito de realizar uma previsão do local em que poderá ocorrer determinado crime dentro de uma janela temporal; bem como identificar indivíduos que estariam envolvidos neste futuro crime, seja como vítima ou como perpetrador dos delitos.

Ainda, os autores, na publicação “Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice” destacam que usualmente os sistemas de policiamento preditivo não são transparentes a respeito de quais dados especificamente são fornecidos aos sistemas para realizar as previsões (Crawford, Richardson; Schultz, 2019)

Todavia, sabe-se que, são primordialmente utilizados dados envolvendo históricos policiais, abrangendo dessa forma, informações referentes a crimes anteriores que se encontram nos registros policiais, englobando a modalidade criminosa, o local e seu horário; bem como dados referentes às abordagens de prisões e chamadas de emergência realizadas para acionar os serviços das autoridades (Crawford; Richardson; Schultz, 2019).

Na mesma esteira, o relatório da Fundação Getúlio Vargas (2023) elucida que o policiamento preditivo trata da combinação de dados previamente conhecidos, seja de fontes convencionais ou não convencionais, que irão produzir estimativas, com uma probabilidade específica, de eventos futuros. E ainda:

> A noção fundamental subjacente à teoria e prática do policiamento preditivo é a possibilidade de inferências probabilísticas sobre atividades criminosas futuras com base em dados existentes. A partir da possibilidade de se utilizar dados de uma ampla variedade de fontes para calcular estimativas sobre fenômenos como onde a violência armada é susceptível de ocorrer, onde um ladrão em série é susceptível de cometer o próximo crime ou mesmo quais indivíduos um suspeito provavelmente contatará para obter ajuda. (Bottino; Fernanda, 2023).

Diante do exposto, é possível extrair que o big data, considerado uma subárea de aplicação da inteligência artificial, vem sendo cada vez mais empregado no contexto da segurança pública. Apesar de seu amplo e controverso conceito, está intimamente ligado a um grande volume de dados e sua análise. Seu uso complementa-se com a inteligência artificial, na medida que esta permite a análise massiva de dados, de forma rápida e eficaz.

Entre seus empregos, destacam-se especialmente as tecnologias de reconhecimento facial, com uma tendência de instalações em espaços públicos de grande circulação para fins de segurança pública; bem como os sistemas de policiamento preditivo, que buscam, com base em bancos de dados de criminalidades prévias, apontar a possibilidade de novos delitos, sua localidade e horário, bem como os indivíduos envolvidos em tais eventos.

Como exposto anteriormente, existem relatórios apontando reduções do índice de criminalidade em locais nos quais tais tecnologias foram empregadas. Todavia, seu uso é controverso e a instalação e emprego de tais sistemas deve ser acompanhada de uma rígida regulamentação e transparência com a população, conforme será exposto no capítulo a seguir.

3. A FALÁCIA DA OBJETIVIDADE NAS DECISÕES DOS SISTEMAS DE INTELIGÊNCIA ARTIFICIAL

Como visto, a segurança pública vem realizando o emprego de novas ferramentas tecnológicas com o intuito de trazer maior eficiência no combate à criminalidade. Todavia, a aplicação traz questionamentos.

Em primeiro lugar, as aplicações de reconhecimento facial em locais públicos levantam debates acerca de sua pertinência e justificabilidade. Fontes e Lutge (2021) questionam o impacto de tais sistemas nos valores democráticos; bem como o impacto na relação entre o indivíduo, o espaço público e o aparelho estatal. Nesse sentido, indagam a aceitabilidade da realização de uma vigilância massiva da população por parte do Estado em prol da ordem e segurança pública.

Outro ponto complexo e controverso é o paradoxo que se cria entre a transparência e a eficiência. Nesse sentido, Fontes e Lutge (2021) apontam que a não comunicação para os cidadãos de que aquele ambiente está sendo vigiado e que sua identidade é rastreada, priva o indivíduo de tomar decisões autônomas sobre a maneira como conduz as atividades de sua vida cotidiana e pública.

Por outro lado, também ressaltam que a divulgação da vigilância de determinado espaço, tende a dispersar as atividades criminosas daquela localidade, apenas realocando-as para outras regiões ainda não monitoradas (Fontes, Lutge, 2021).

Também de suma importância é a discussão envolvendo a veracidade e objetividade dos resultados apontados pelos sistemas. Existe uma falsa percepção de que as decisões apontadas pelos sistemas são revestidas de objetividade e fidedignidade, tendo em vista que se tratam de máquinas. Todavia, é preciso compreender que os resultados dos sistemas estão diretamente relacionados à programação dos algoritmos e aos dados que lhe foram fornecidos, sendo que tais fatores possuem interferência humana, sendo, dessa forma, passíveis de subjetividade e reprodução de estigmas sociais.

Dentro dessa lógica, notável o estudo realizado por Crawford, Richardson e Schultz (2019), cujos resultados foram publicados no artigo “Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice”. Referida pesquisa buscou investigar treze jurisdições que desenvolveram e implementaram sistemas de policiamento preditivo ao mesmo tempo em que estavam sendo investigadas por práticas de policiamento corruptas, especialmente racistas e ilegais.

O objetivo do estudo recai em questionar os dados utilizados para alimentar os sistemas de policiamento preditivo. Apontaram a existência de departamentos policiais que possuíram comprovadamente práticas ilegais, corruptas e racistas, que abrangiam desde abordagens policiais até adulteração de registros oficiais. Tendo em vista que os sistemas são alimentados primordialmente por dados de históricos policiais, levanta-se a preocupação de que dados sujos levarão à criação de sistemas enviesados.

Logo, cria-se o que os autores chamam de "feedback loop". Ou seja, é estabelecido o risco de que, determinados indivíduos, de certas categorias marginalizadas (imigrantes, negros, latinos) sejam apontados pelo sistema como mais propensos a cometer crimes, pois os indivíduos dessas categorias são abordados por policiais com uma maior frequência. E o sistema, configurado com base nesses dados enviesados, apenas os reproduzirá, perpetuando discriminações.

A mesma lógica se aplica para predição em relação à locais. Determinadas regiões da cidade são alvo de maior atenção policial, logo, o sistema apontará tais locais como regiões com maior probabilidade de ocorrer novas práticas delitivas. A situação se agrava, pois os sistemas transmitem uma falsa percepção de fidedignidade e objetividade.

Como exposto no capítulo anterior, o policiamento preditivo envolve sistemas que realizam a análise de dados com o intuito de prever onde um crime pode ocorrer, bem como quais são as pessoas envolvidas no delito, seja na condição de vítima ou perpetrador do crime. Os autores apontam para a baixa transparência dos fornecedores do sistema, pois não especificam quais dados são utilizados, nem apontam de que forma ocorrerá a responsabilização em caso de imprecisões ou decisões incorretas (Crawford, Richardson e Schultz, 2019).

Apesar da falta de transparência, sabe-se, de maneira geral, que os dados utilizados pelos sistemas são, principalmente, registros policiais, englobando informações como tipo criminal, o horário nos quais os delitos foram realizados, bem como o local e ligações telefônicas de denúncias para atendimento (Crawford; Richardson; Schultz, 2019).

Todavia, é possível que essas informações criem os chamados "dados sujos". Primordialmente, é importante esclarecer o significado e abrangência deste termo empregados pelos autores. "Dirty data" ou "dados sujos" é uma expressão utilizada para se referir a ausência de dados; dados incorretos ou representações não padrão dos mesmos dados. Nesse sentido, na pesquisa em tela, refere-se a dados criados a partir de práticas corruptas, enviesadas ou ilegais; bem como informações que tenham sido intencionalmente manipuladas (Crawford; Richardson;Schultz, 2019).

Como exemplo é possível apontar registros policiais de prisões de pessoas inocentes, falsamente acusadas, especialmente decorrente de abordagens e prisões com viés racial; assim como evidências plantadas. Ainda, o estudo aponta para a existência recorrente de manipulação dos registros policiais com o intuito de fraudar as estatísticas criminais, por motivos políticos, de financiamentos ou para promoção de relações públicas (Crawford; Richardson ; Schultz, 2019).

Desse modo, cria-se uma preocupação em relação a como esses dados sujos são tratados pelos sistemas de policiamento preditivo. Crawford, Richardson e Schultz (2019) apontam que mesmo fornecedores que reconhecem a existência desses dados, não apresentam uma solução satisfatória, limitando-se a isolar e excluir determinadas informações. Todavia, a mera existência desses dados enviesados seria suficiente para questionar a fidedignidade de todo o conjunto de informações de uma jurisdição.

A PredPol, fornecedora de sistemas de policiamento preditivo, buscando esquivar-se de dados sujos, afirma, por exemplo, que realiza a exclusão de informações referentes a delitos que possuam relação com entorpecentes, bem como abordagens de trânsito, tendo em vista a reconhecida existência de disparidades raciais nessas situações. Todavia, apesar do esforço da fornecedora, a discricionariedade de autoridades policiais não se limita apenas a essas ocorrências, dessa forma, o risco subsiste (Crawford, Richardson; Schultz, 2019).

Buscando embasar as preocupações de seus estudos, Crawford, Richardson e Schultz (2019) realizaram um estudo de caso de três jurisdições, sendo elas: Chicago, na qual, comprovadamente o sistema preditivo foi alimentado com “dados sujos”; Nova Orleans, onde apontam um risco extremamente alto de que os “dados sujos” tenham sido utilizados no sistema de policiamento preditivo e o Condado de Maricopa, local com expressivas evidências da existência de criação de “dados sujos”, e cujo sistema de policiamento preditivo falha na transparência pública.

Para os fins exemplificativos dessa pesquisa, será analisado o caso de Chicago. No ano de 1972, um Comitê Blue-Ribbon[1] apontou a existência de práticas policiais racistas. Ato contínuo, as investigações apontaram para evidências de tortura de mais de cem homens negros entre os anos de 1972 a 1991, bem como a existência de um processo judicial que apontava o envio desproporcional de viaturas e chamadas de emergência para bairros cujas populações eram especialmente compostas por minorias étnicas (Crawford; Richardson; Schultz, 2019).

Somado a isso, no ano de 2015, a organização não governamental norte-americana ACLU (União Americana pelas Liberdades Civis) emitiu um relatório denunciando as

abordagens realizadas pelas autoridades policiais de Chicago. Referido relatório constatava a existência de abordagens ilegais, bem como um desproporcional número de abordagens de indivíduos negros (Crawford, Richardson; Schultz, 2019).

A morte do jovem Laquan McDonald por um oficial, levou o Departamento de Justiça a uma investigação do Departamento de Polícia de Chicago (CPD). Como resultado, o relatório apontou que:

> O Departamento de Policia de Chicago estava envolvido num padrão ou prática de uso inconstitucional da força; má recolha de dados para identificar e abordar condutas ilegais; deficiências sistêmicas na formação e supervisão; deficiências sistêmicas nos sistemas de responsabilização que contribuem para o padrão ou prática de conduta inconstitucional; e conduta inconstitucional que afeta desproporcionalmente residentes negros e latinos (Crawford, Richardson; Schultz, 2019) (tradução livre).

Enquanto tais investigações e relatórios apontavam as práticas desviantes do Departamento, foi desenvolvida a Strategic Subject List (SSL). Tal ferramenta foi criada pelo Instituto de Tecnologia de Illinois e sua implementação data do ano de 2012. A ferramenta incorporava dados e buscava classificar e identificar indivíduos que poderiam se tornar vítimas, ou possíveis perpetradores de tiroteios, ou homicídios (Crawford; Richardson; Schultz, 2019).

Posteriormente, no ano de 2017, com a Lei de Liberdade de Informação, fatos sobre a ferramenta se tornaram públicos, entre eles o conhecimento que o conjunto de dados do sistema incluía 398.684 indivíduos (Crawford, Richardson; Schultz, 2019).

Ainda, sabe-se que a ferramenta trabalha através da classificação de indivíduos em níveis “muito baixo” até “muito alto”, ou seja, expressando a probabilidade da pessoa ser vítima ou perpetradora do delito. Somado a isso, os dados levados em consideração pelo sistema são majoritariamente baseados em registros de abordagens e prisões, ou seja, realizados pelo departamento de polícia, sendo eles:

> O número de vezes que um indivíduo foi vítima de tiroteio; a idade do indivíduo durante a última prisão; o número de vezes que o indivíduo foi vítima de agressão ou agressão agravada; tendências na atividade criminosa; o número de prisões anteriores por uso ilegal de arma; o número de prisões anteriores por crimes violentos; o número de prisões anteriores por narcóticos; e afiliação a gangues (Crawford, Richardson; Schultz, 2019) (tradução livre).

Nota-se uma maior tendência a informações baseadas nos registros policiais em detrimento de dados a respeito de indivíduos que foram efetivamente condenados. Nessa esteira, um estudo do The New York Times, em conjunto com a Upturn apontou que um terço dos indivíduos na lista jamais foram efetivamente presos ou vítimas de crime. E ainda,

aproximadamente 70% da lista foi classificada com uma pontuação de alto risco. Por fim, os dados indicam que 56% de todos os homens negros com menos de 30 anos de idade em Chicago possuem uma pontuação de alto risco.

Do exposto, depreende-se que a aplicação de algumas tecnologias de inteligência artificial na segurança pública tais como o reconhecimento facial e o policiamento preditivo levantam questões acerca da objetividade dessas ferramentas. Apesar de dados apontarem uma possível eficiência na redução de índices de criminalidade pelos sistemas, também destacam-se alguns dilemas éticos como transparência com a população e a objetividade dos dados utilizados. A análise realizada por Crawford, Richardson e Schultz (2019) revela a criação de ferramentas de policiamento preditivo alimentadas com registros policiais de jurisdições investigadas e denunciadas por corrupção, alteração de registros e racismo e o emprego desses dados na concepção do sistema pode levar a perpetuação de discriminações.

4. PERSPECTIVAS ÉTICAS E NORMATIVAS DO USO DA INTELIGÊNCIA ARTIFICIAL NA SEGURANÇA PÚBLICA

Como exposto, os sistemas de inteligência artificial tendem a passar uma falsa percepção de que seus resultados são objetivos, fidedignos e imparciais. Todavia, sabe-se que seus resultados estão diretamente ligados à sua programação, bem como aos dados utilizados durante o seu treinamento.

Nessa mesma esteira, a Fundação Getúlio Vargas (2020) aponta que a existência de sistemas autônomos cria uma aura de objetividade, especialmente em razão de seus processos complexos. Todavia, características subjetivas e julgamentos humanos podem ser inseridos nesses sistemas através do design dos algoritmos, na medida em que é neste momento que são definidas as características da ferramenta; são delimitados os dados para o seu treinamento e definidos seus limites e parâmetros (Abbas; *et al*, 2020).

Diante desse cenário, é imprescindível a discussão de diretrizes para sua implementação e uso, norteadas por princípios éticos, buscando coibir o uso indevido dessa tecnologia. Nesse mesmo sentido:

> A IA demanda reflexões éticas contemporâneas justamente por provocar situações limites que vão além da velocidade de processamento, capacidade de armazenamento, busca de informações, padronizações típicas de sistemas de automação, mas especialmente por dois fatores: a execução de atividades cognitivas fruto de sistemas de aprendizagem de máquina ou, no mínimo, a delimitação do conteúdo (diante do volume astronômico de dados de nossos tempos) sobre o qual a cognição humana irá atuar para promover suas decisões [...]. (Peixoto, 2020)

Em resposta à consulta pública do Ministério da Ciência Tecnologia e Comunicações-MCTIC sobre a Estratégia Brasileira de Inteligência Artificial, a Fundação Getúlio Vargas elaborou o "Policy Paper: Regulação de Inteligência Artificial no Brasil", no qual apresenta algumas diretrizes para a normatização do emprego desses sistemas.

Inicia-se apontando alguns princípios éticos que devem ser seguidos para o adequado uso dessas tecnologias. A priori menciona diretrizes elencadas pela Organização para a Cooperação e Desenvolvimento Econômico (OCDE), tendo em vista que o Estado brasileiro aderiu a esses princípios, portanto, devem ser levados em consideração (Abbas; *et al*, 2020). Entre eles destacam-se:

> i) promover o crescimento inclusivo, o desenvolvimento sustentável e o bem-estar;
> ii) respeitar o Estado de Direito, os direitos humanos, os valores democráticos e a diversidade, liberdade, dignidade, autonomia, privacidade e proteção de dados, não-discriminação e igualdade, diversidade, equidade, justiça social e direitos trabalhistas internacionalmente reconhecidos. Para isso, devem incluir salvaguardas apropriadas, como possibilitar a intervenção humana sempre que necessário;
> iii) deve haver transparência e explicabilidade sobre os sistemas de IA.

Somado a isso, a Fundação, através de suas pesquisas, identificou outros princípios que também deveriam nortear a normatização brasileira como, por exemplo, a privacidade, abrangendo a ideia de que os indivíduos devem consentir com a implementação desses sistemas; assim como a necessidade da existência de uma transparência no que tange o uso e processamento dos dados e o direito à retificação (Abbas; *et al*, 2020).

Também é ressaltada a importância de serem observadas a equidade e a não-discriminação; assim como a necessidade de se existir um controle humano sobre as tecnologias, este último estando diretamente relacionado com outra diretriz: o direito de uma revisão humana da decisão automatizada (Abbas; *et al*, 2020).

Ressalta-se neste ponto, que os princípios norteadores devem ser aplicados em todas as fases de um sistema de inteligência artificial. Assim, devem estar presentes durante a concepção e criação da ferramenta, ou seja, no momento em que são definidos seus limites e parâmetros, bem como os dados fornecidos para seu treinamento(Abbas; *et al*, 2020).

As indicações éticas também devem estar presentes durante a implementação do sistema na sociedade, na medida em que a entidade implementadora se certifique que o produto adquirido cumpre com os requisitos éticos desejados. E ainda, é importante que a sociedade seja informada acerca desses requisitos e atue exigindo que sejam respeitados e observados (Abbas; *et al*, 2020).

Aponta-se que a incorporação dos princípios na arquitetura dos sistemas de inteligência artificial pode-se apresentar de duas formas: através de abordagens bottom-up (“de baixo para cima”) ou top-down (“de cima para baixo”).

A FGV esclarece que a abordagem bottom-up não é a ideal para a aplicação de princípios. Nesse caso, de forma sintetizada, o sistema de inteligência artificial observa o comportamento humano e aprende a tomar decisões com base nas condutas identificadas. Todavia, destaca-se o caso do chatbot Tay, da Microsoft, o qual, ao interagir com os usuários do X (antigo Twitter) começou a adotar discursos racistas e agressivos politicamente, sendo tais discursos ensinados e influenciados pelos próprios usuários através da interação com o sistema (Abbas; *et al*, 2020).

Por outro lado, a abordagem top-down mostra-se mais adequada para a implementação de princípios éticos, nesse modelo, as regras e os princípios são programados no sistema. Para tanto, é possível que seja adotada uma “white list”, ou seja um conglomerado de regras que o sistema sempre deverá observar durante o seu funcionamento; bem como uma “black list”, traduzidas como comportamentos que o sistema sempre deverá evitar (Abbas; *et al*, 2020). Cabe ressaltar nesse momento, a observação realizada pela FGV em relação aos algoritmos de aprendizagem de máquina:

> Algoritmos mais complexos, que conseguem adaptar dinamicamente o seu comportamento, e, por isso, podem apresentar um comportamento inesperado. Sendo assim, sistemas de machine learning exigem que os princípios e regras sejam integrados nas três etapas do ciclo: i) na etapa de “percepção”, o sistema deve ser desenvolvido de modo a reconhecer todos os elementos ambientais necessários para assegurar que os princípios sejam respeitados; ii) na etapa de “planejamento”, o sistema apenas deve considerar os planos que cumprirem com os requisitos; iii) na etapa de “ação”, as ações do sistema devem restringir-se aos comportamentos que cumprem os requisitos (Abbas; *et al*, 2020).

Por fim, em seu relatório a FGV (2020) destaca a importância da existência de métodos de explicabilidade e transparência nesses sistemas. Nesse sentido informa que a explicabilidade está diretamente relacionada à ideia de que seja possível expor de forma lógica e fundamentada os motivos que levaram a determinada decisão pelo sistema. E complementa: “é importante que um observador externo possa compreender em que medida certo fator exerceu influência ou foi determinante para o resultado”. Convém ressaltar que o direito à explicação dos fundamentos de decisões automatizadas para os indivíduos afetados já é previsto na Lei Geral de Proteção de Dado(Abbas; *et al*, 2020).

Ressalta-se nesse ponto que os sistemas de machine learning demandam uma maior atenção, na medida em que, em razão da sua natureza, o funcionamento desses sistemas pode

ir além dos limites da compreensão humana. É possível a existência de algoritmos que aprendam sozinhos, através de sua experiência prática. Ou seja, nesse cenário, serão incorporados novos dados além daqueles apresentados em seu treinamento (Abbas; *et al*, 2020).

Dessa maneira, seria possível que algumas decisões tomadas por esses algoritmos não sejam facilmente explicadas, pois levam em consideração dados inseridos além de seu treinamento. Por cautela, o relatório sugere que determinados algoritmos de machine learning sejam evitados em áreas mais críticas, ou seja, que demandem uma explicação. Nesse cenário, o policiamento preditivo e o reconhecimento facial poderiam apresentar tais dificuldades.

Não obstante, a Fundação aponta a existência de pesquisas que buscam desenvolver algoritmos de machine learning que sejam capazes de aplicar raciocínio causal ou contrafactual. Dessa maneira, haveria a possibilidade de apresentar uma cadeia lógica de informações de causalidade para julgar a tomada de decisão do algoritmo, sendo portanto, uma possível solução para o problema apresentado.

5. CONCLUSÃO

Os recentes avanços tecnológicos geram novos debates acadêmicos e dilemas éticos. A inteligência artificial vêm sendo amplamente discutida em razão de suas diversas aplicações. Não existe um consenso acerca do conceito de inteligência artificial, todavia, sabe-se que esta está ligada a integração de sistemas computacionais e algoritmos para a realização de tarefas cognitivas.

Assim, em decorrência dos motivos expostos no presente trabalho, nota-se ser necessário o estabelecimento de uma legislação mais ampla, visando fornecer maior suporte ao judiciário e aos tribunais na tomada de decisões voltadas à inteligência artificial. Além disso, também são importantes normas que estabeleçam formas claras de limitação do uso, ou ainda, capazes de determinar que as empresas que programam fazem o uso de algoritmos e sistemas a base de inteligência artificial sejam mais transparentes.

Contudo, ainda que a solução pareça relativamente simples, ela vem acompanhada de um dilema: dado o contexto de constante expansão tecnológica, uma legislação muito específica, que abarque todos os aspectos em torno da IA, pode se tornar obsoleta de forma mais rápida do que se imagina. Ante ao exposto, em decorrência dessa nova realidade tecnossocial, resta evidente a importância em se realizar debates voltados ao trâmite de

regulamentações e propostas legislativas que visam reduzir os impactos negativos da IA no Brasil.

Do exposto, depreende-se que a aplicação de algumas tecnologias de inteligência artificial na segurança pública devem ser realizadas com cautela, em razão das questões envolvendo a objetividade dessas ferramentas. Os sistemas de IA refletem uma falsa percepção de neutralidade, contudo, seus resultados estão diretamente relacionados com sua programação. A Fundação Getúlio Vargas, em seu relatório para a regulamentação de IA, ventila algumas importantes questões éticas que deveriam ser observadas para o devido uso dessas ferramentas, destacando-se a transparência e explicabilidade.

REFERÊNCIAS

ABBAS, Lorena; *et al.* **Policy Paper:** Regulação da inteligência artificial no Brasil. Rio de Janeiro: FGV Direito Rio, 2020.

ANDRADE, N. G de; DONEDA, D.C.M.; MENDES, L.S; SOUZA, C.A.P. de. Considerações iniciais sobre inteligência artificial, ética e autonomia pessoal. **Pensar, Fortaleza,** v. 23, n. 4, p. 1-17, out./dez. 2018. DOI: 10.5020/2317-2150.2018.8257. Disponível em: https://ojs.unifor.br/rpen/article/view/8257. Acesso em: 1 maio 2024.

BOTTINO, Thiago; FERNANDA, Daniel Vargas (coord). **Segurança pública na era do big data.** Rio de Janeiro: FGV Direito Rio, 2023.

COSTA, Ramon Silva; KREMER, Bianca. Inteligência artificial e discriminação: desafios e perspectivas para a proteção de grupos vulneráveis frente às tecnologias de reconhecimento facial. **Revista Brasileira de Direitos Fundamentais & Justiça**, *[S. l.]*, v. 16, n. 1, 2022. DOI: 10.30899/dfj.v16i1.1316. Disponível em: https://dfj.emnuvens.com.br/dfj/article/view/1316. Acesso em: 28 maio. 2024.

CRAWFORD, Kate; RICHARDSON, Rashida; SCHULTZ; Jason M. Dirty Data, Bad Predictions: How civil rights violations impact police data, predictive policing systems and justice. **New York University Law Review** *[S.l.],* v. 94, n. 15, 2019. Disponível em:https://www.nyulawreview.org/online-features/dirty-data-bad-predictions-how-civil-rights-violations-impact-police-data-predictive-policing-systems-and-justice/. Acesso em: 26 maio 2024.

FONTES, Ana Catarina; LÜTGE, Christoph. Vigilância e Relações de Poder – O Uso de Tecnologias de Reconhecimento Facial e Identificação Biométrica a Distância em Espaço Público e Impactos na Vida Pública. **Revista Direito Público**, Brasília, Vol. 18, n. 100, p. 81-106, out./dez. 2021. Disponível em: https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/6203. Acesso em: 22 maio 2024.

KAUFMAN, Dora. **A inteligência artificial irá suplantar a inteligência humana?**. – 1ª. ed; – Barueri–SP: Estação das Letras e Cores, 2019.

KAUFMAN, Dora. **Desmistificando a inteligência artificial**. – 1ª. ed; – Belo Horizonte: Autêntica, 2022.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS PARA A EDUCAÇÃO, A CIÊNCIA E A CULTURA (UNESCO). Representação da UNESCO no Brasil. *Recomendação sobre a Ética da Inteligência Artificial.* 2022. Disponível em: https://unesdoc.unesco.org/ark:/48223/pf0000381137_por. . Acesso em: 1 maio. 2024.

PEIXOTO, Fabiano Hartmann. **Inteligência artificial e direito:** convergência ética e estratégica - 1.ed. – Curitiba: Alteridade Editora, 2020.

RIGANO, Christopher. Using Artificial Intelligence to Address Criminal Justice Needs. **National Institute of Justice Journal.** Washington, Vol. 280, p. 36-46, Jan. 2019. Disponível em: https://nij.ojp.gov/nij-journal/nij-journal-issue-280. Acesso em: 22 mai 2024.

RUSSELL, Stuart J.; NORVIG, Peter. **Inteligência Artificial: Uma Abordagem Moderna.** Tradução Daniel Vieira; Flávio Soares Corrêa da Silva. - 4. ed. - Rio de Janeiro: Grupo Editorial Nacional S.A, 2022.

NOTAS TEXTUAIS

[1] Grupo de pessoas designado para realizar investigação e análise de determinadas situações.