



INTELIGÊNCIA ARTIFICIAL NO DIREITO: DILEMAS E CONTRIBUIÇÕES ARTIFICIAL INTELLIGENCE IN LAW: DILEMMAS AND CONTRIBUTIONS

Flávio Iassuo Takakura¹

Luciana Gaspar Melquíades Duarte²

RESUMO

Um dos mais importantes aspectos da Inteligência Artificial (IA) é a inevitabilidade da presença de vieses. Neste trabalho, apresenta-se uma pesquisa qualitativa apontando as possíveis causas, efeitos e malefícios desses vieses e como mitigá-los. Uma das formas de mitigação é a democratização do processo de criação e execução da IA, que, contudo, em determinados tipos de IA, como na *Deep Learning*, não se revela suficiente, uma vez que o processo decisório é opaco a humanos. No âmbito do Direito, devido aos vieses, a IA deve ser utilizada em estreita participação com humanos, devendo caber, sempre, a decisão final ao homem.

PALAVRAS-CHAVE: Inteligência artificial; vieses; mitigação; Direito; decisão final.

ABSTRACT

One of the most important aspects of Artificial Intelligence (AI) is the inevitability of the presence of biases. In this work, a qualitative research is presented, pointing out the possible causes, effects and harm of these biases and how to mitigate them. One way to mitigate is the democratization of the all process, creation and execution, which, however, in certain types of AI, such as Deep Learning, is not enough, since the decision-making process is opaque to humans. In Law, due to biases, AI must be used together with humans, and the final decision must always be up to human.

¹ Mestrando em Direito e Inovação da Faculdade Direito da UFJF. Mestre e Doutor em Física pela Pontifícia Universidade Católica do Rio de Janeiro. Professor Titular da Universidade Federal de Juiz de Fora. Endereço: Universidade Federal de Juiz de Fora, Instituto de Ciências Exatas – Departamento de Física, s/n, Cidade Universitária 36013-020 - Juiz de Fora, MG – Brasil. Endereço eletrônico: flavio.takakura@ufjf.edu.br.

² Mestre e Doutora em Direito pela Universidade Federal de Minas Gerais. Professora Adjunta de Direito Constitucional e Administrativo da Universidade Federal de Juiz de Fora, Pesquisadora, Professora do Mestrado em Direito e Inovação da Faculdade de Direito da UFJF e ex-Procuradora do Município de Juiz de Fora. Endereço: Universidade Federal de Juiz de Fora, Faculdade de Direito, s/n, Cidade Universitária 36013-020 - Juiz de Fora, MG – Brasil. Endereço eletrônico: lg.melquiades@uol.com.br.





KEY-WORDS: *Artificial intelligence; biases; mitigation; Law; final decision*





1 INTRODUÇÃO

Muito se fala em Inteligência Artificial (IA) (ORACLE, 2021) nos dias atuais, devido ao seu uso como assistentes nos aparelhos de *smartphones*, como atendentes virtuais nos aplicativos de lojas e bancos, ou ainda mais remotamente, por conta de filmes de ficção científica. Questiona-se, contudo, se a IA pode vir a substituir a inteligência humana e, assim, representar uma ameaça à humanidade.

Os temores éticos e morais inerentes a este questionamento são usualmente decorrentes da possível introdução de vieses na programação eletrônica. Indaga-se se eles poderiam fazer com que a IA traga, ética e moralmente, consequências inaceitáveis para os padrões da sociedade.

Diante dessas questões, o uso da IA no Direito suscita precauções, considerando a potencialidade de influência indevida dos vieses nas decisões jurídicas.

Objetivando evitar que os eventuais vieses que surjam, possam ser inseridos ou sejam exacerbados na operacionalização de sistemas que utilizem IA, e em que grau, propiciar resultados que afrontem os valores da sociedade, esta pesquisa foi efetuada em caráter exploratório, através de pesquisa bibliográfica sobre aspectos técnicos de programação de IA e do ponto de vista da Filosofia, Sociologia e do Direito (CANZIANI, 2021).

2 DEFINIÇÕES

O dicionário *Michaelis on line* (MICHAELIS, 2021) define viés como o “processo de distorção da percepção da realidade, como resultado do envolvimento do observador naquilo (sociedade, cultura, classe etc.) que está observando”. Outro significado também atribuído pelo mesmo dicionário é o do “erro sistemático ou tendenciosidade”. Portanto, no contexto da IA, o viés decorre de um erro sistemático que pode ocorrer por diversos fatores, como a necessidade de se ter algoritmos que possam produzir resultados/previsões em um intervalo de tempo curto com o baixo consumo de tempo de computação. Isso leva a produção de algoritmos que sejam o menos complexo possíveis, ou seja, tais que não levem em conta todas as nuances da realidade que se pretende analisar.

O conceito de IA tem origem (PRADO, 2016; WIKIPEDIA, 2021) nos primórdios do conhecimento. Na Grécia Antiga, de acordo com Marar (PRADO, 2016), os filósofos gregos



como Platão e Aristóteles idealizavam uma inteligência não humana que pensasse por si. O que há de revolucionário nisso é a linha de pensamento: eles já idealizavam a ciência cognitiva, que trata da aprendizagem humana.

Entre várias definições possíveis, a IA pode ser considerada como o ramo da ciência da computação que se propõe a elaborar sistemas computacionais - artificiais - que simulem a inteligência humana, indo além da programação de ordens específicas, com a capacidade de raciocinar, perceber, resolver problemas e tomar decisões de forma autônoma, baseadas em padrões obtidos de enormes bases de dados³.

Essa capacidade de tomada de decisão de forma autônoma só é possível se a máquina for capaz de raciocinar. Entretanto, pensar é uma habilidade apenas da raça humana, que, contudo, desenvolveu ferramentas capazes de viabilizar que uma máquina pense. Do ponto de vista estritamente técnico, isto se dá mediante a introdução de um programa - *software* - elaborado por um programador humano, na máquina física - *hardware*, capaz de fazer com que o *hardware* execute os comandos contidos no *software*. Assim, tudo que a máquina faz depende do *software*.

A capacidade de pensar está associada umbilicalmente com a consciência, que seria uma característica inerente aos seres vivos, segundo autores como Quaresma (QUARESMA, 2019) e Searle (SEARLE, 1997), que defendem um modelo monista em contraponto à proposta de Descartes e Galileu (KANDEL, 2021) formulada no século XVII, da existência de uma distinção entre a realidade física descrita pela ciência e a realidade mental da alma, que estaria fora do escopo da pesquisa científica, ou seja, um dualismo entre uma matéria consciente e uma inconsciente. Quaresma (2019) e Searle (1997) defendem que inexiste tal dualismo e que a consciência é um fenômeno biológico ordinário como crescimento, digestão, etc. Para autores monistas, a consciência não é uma causa dos processos que ocorrem no cérebro, pois, se assim fosse, seria um efeito causal e poder-se-ia dizer que seria dualista. A consciência seria, isto sim, característica do cérebro e não um efeito conectado por uma causalidade.

Os defensores do monismo vão além e consideram a intencionalidade, assim como a consciência, como uma característica do cérebro e, logo, uma vez que o resultado de qualquer

³ Bases de dados são compreendidas como fontes de informação eletrônicas contendo conjunto de dados inter-relacionados, pesquisáveis de modo interativo ou conversacional através de um computador para permitir a recuperação da informação (BIBLIOTECONOMIADIGITAL, 2017).



processo de solução de problema é intencional, teria que ser, em última instância, um resultado de um processo que envolva necessariamente um ser biológico, visto que, se a intencionalidade é uma característica do cérebro, seria, portanto, um fenômeno biológico (QUARESMA, 2020). Para estes autores, também a emoção seria característica do cérebro, e, assim, concluem pela impossibilidade de atribuir inteligência a seres não biológicos, o que leva os defensores do monismo a rejeitarem, do ponto de vista do Direito, a autonomia da IA levantada em discussões sobre a atribuição de “personalidade eletrônica” aos robôs com IA como a materializada na Resolução do Parlamento Europeu (PARLAMENTO EUROPEU, 2017). Se centrada ou não no antropocentrismo - o antropomorfismo, como metáfora para caracterizar a IA ou robôs com IA, ou em analogia à atribuição da personalidade jurídica (NEGRI, 2020) – há um movimento no Direito no sentido de abertura para novas formas de subjetividade, em que não há relevância sobre o modelo adotado, se monista ou se dualista, sendo, em ambos, inegável a existência de vieses.

Investiga-se qual a diferença do software utilizado na IA com os demais utilizados rotineiramente. Na IA, o software é elaborado para, ao analisar um grande grupo de dados, identificar padrões que são armazenados, podendo ou não ser comparados com um padrão previamente dado, ou seja, quando novos dados são inseridos, a IA compara esses novos dados aos armazenados e faz uma previsão ou toma uma decisão. Se sua previsão/decisão estiver correta, ela armazena como acerto; se, ao revés, estiver errada, armazena a informação de erro e assim continua procedendo, de forma que, quanto mais dados for recebendo, “melhor” se torna seu padrão de referência e cada vez mais, o índice de acerto aumenta, fazendo com que seja possível à IA fazer previsões ou tomar decisões, que caracteriza o processo de pensar dela.

Já os algoritmos utilizados na programação rotineira, tipo não IA, caracterizam-se por serem compostos de conjuntos finitos de comandos ou passos a serem executados para a solução de um determinado problema. Assim, cada passo a ser executado é claramente definido por meio do algoritmo executado.

A IA é o grande grupo de sistemas computacionais que se propõe a simular a inteligência humana e que contém um subgrupo, o *machine learning* - aprendizado de máquina, em que se encontram os algoritmos que executam essa tarefa de pensar. Esse subgrupo pode ser classificado como sendo supervisionado, não supervisionado ou aprendizado por reforço.



No aprendizado de máquina supervisionado, é fornecido um conjunto de dados iniciais para treinar a máquina para que ela, com base neles, execute tarefas para atingir um determinado resultado previamente obtido. Após esse treinamento inicial, outro conjunto de resultados consolidados é confrontado com os resultados obtidos pelo *machine learning*, por isso, supervisionado. Caso não haja uma coincidência entre os resultados consolidados e os obtidos pela IA, é necessário continuar a treiná-la, oferecendo-lhe mais dados, até que atinja o ponto considerado ideal para que possa entrar em operação e a partir de dados inéditos, produzir resultados/previsões.

O aprendizado não supervisionado é aquele no qual o conjunto de dados não é previamente classificado e a IA deve, com base nesses dados iniciais, classificá-los.

No aprendizado por reforço, o algoritmo é construído para que o sistema aprenda por tentativa e erro, ou seja, toda vez que há um acerto com vistas a maximizar o sucesso, há uma recompensa. Não há interferência humana e nem uma base de dados. Exemplo desse tipo de IA é o processo de aprendizado de andar de uma IA. Para todas essas classificações utilizam-se, geralmente, as redes neurais (artificiais).

Segundo Carvalho (CARVALHO, [s.d]), “Redes Neurais Artificiais são técnicas computacionais que apresentam um modelo matemático inspirado na estrutura neural de organismos inteligentes e que adquirem conhecimento através da experiência.”

Finalmente, dentro do subgrupo de *machine learning*, há um outro grupo menor, conhecido como *deep learning* ou aprendizado profundo. *Deep Learning* difere do *machine learning* porque, nesse, o modelo simula uma rede neural de uma camada, enquanto, naquele, há mais camadas de neurônios.

A existência de mais camadas gera uma grande alteração em relação ao modelo de uma camada. No de uma camada, *machine learning*, dado um conjunto de dados, é possível inferir quais deram origem ao resultado, ou seja, o algoritmo é transparente – se não para o público em geral, ao menos para os especialistas. Já no caso do *deep learning*, isso não ocorre. O algoritmo é opaco, pois, com a introdução de mais camadas de “neurônios”, quanto mais camadas se tem, mais opaca se torna, no sentido de que dado um resultado, não é, em tese, possível saber quais caminhos foram percorridos pelos dados que deram origem ao mesmo.

A partir da apresentação acima, nota-se que é o *software* que faz com que a IA siga um ou outro caminho, determinando, inclusive, a qualidade das projeções ou decisões. Portanto, o

momento da elaboração do software é um dos que possibilita a inserção dos vieses (GOMES, 2019; MANYIKA, 2019; PROGRA{M}ARIA, 2020; VIEIRA, 2019) na IA.

Outra possibilidade de inserção dos vieses é por meio dos dados utilizados, tanto os que definem o padrão inicial como os inseridos posteriormente. Se o padrão inicial já apresentar um viés, os demais dados a serem comparados com ele serão considerados acertos se com eles apresentarem semelhanças, fazendo-se necessária uma intervenção externa para correção a fim de evitar a perpetuação e/ou ampliação do viés inicial. Também quando os novos dados apresentam vieses, mesmo que o padrão inicial não os apresente, pelo processo de aprendizagem, eles podem vir a ser acentuados, de forma que, novamente, faz-se necessária uma intervenção externa para a sua correção.

3 VIESES

Há diversas espécies de vieses (PROGRA{M}ARIA, 2020); porém, aquelas que influem decisivamente na IA podem ser geralmente classificadas em quatro tipos, quais sejam, os vieses de algoritmo, de amostra, de medida e de cognição.

Para a compreensão do viés de algoritmo (*Algorithm Bias*), necessário enfrentar, primeiramente, o conceito de algoritmo. Baer (BAER, 2019) os define como “equações matemáticas ou outras regras lógicas para resolver um problema específico”⁴ e, portanto, o viés algorítmico surge porque os algoritmos utilizados na IA procuram solucionar um determinado problema simulando o processo que o cérebro humano utiliza para resolver problemas que demandem soluções rápidas. Por exemplo, quando uma pessoa, ao atravessar a rua, percebe um evento inesperado como um carro indo de encontro a ela. Essa pessoa não faz todo o raciocínio lógico calculando a velocidade do carro, o tempo que ele levará para atingi-la, mas sim, usará de um processo de “cortar caminho” ou seguir atalhos nesse processo para tomar sua decisão; logo, a esquivada ocorrerá sem a completa análise dos fatos, podendo advir uma decisão equivocada, tomada pelo sistema cognitivo 1 (pensamento rápido, intuitivo e emocional), sendo, portanto, mais propenso a erros em oposição ao outro sistema, cognitivo 2, mais metódico, lógico, ponderado e também mais lento, mas com menos possibilidades de cometer erro (KAHNEMAN, 2011; ONUMULHERES, 2016). Como nesse exemplo, os

⁴ Baer, Tobias. *Understand, Manage, and Prevent Algorithmic Bias*. Apress, Berkeley, Ca. Tradução livre. Texto original: “*Algorithms are mathematical equations or other logical rules to solve a specific problem*”.



algoritmos são construídos para entregar soluções no mais breve intervalo de tempo possível, com menor gasto de tempo de computação, fazendo com que surja, intrinsecamente, a possibilidade de errar. Esse viés, em particular, nada tem a ver com os dados, referindo-se a uma propriedade do próprio algoritmo de IA. Nesse contexto, o viés descreve uma propriedade do algoritmo que influencia o desempenho da previsão/decisão.

O viés de amostra⁵ ocorre quando os dados utilizados para treinar o modelo não representam com precisão o universo/população sobre o qual a IA atuará. Exemplo disso é o da falta de representatividade da amostra, que não contempla toda a variabilidade da população, ou seja, certos “indivíduos” tem mais chance de serem selecionados em uma amostra aparentemente representativa do universo em análise, superestimando ou subestimando a probabilidade de ocorrência do evento no universo real (VIEIRA, 2019; COWGILL, 2021; BBC, 2015). Muitas vezes, esse viés ocorre por conta da escolha inadequada do tipo de amostragem. Observe, ainda, que, quanto maior for o tamanho do universo, mais sensível é o procedimento estatístico ao viés da amostra e, assim, quanto maior for o viés, pior se torna o resultado (MORETTIN, 2010).

O viés de medida decorre de erro na alimentação dos dados provenientes de medições equivocadas efetuadas por um dispositivo. Por exemplo, no caso de um programa de reconhecimento de cores, se o dispositivo utilizado para a captura das informações (uma câmera, por exemplo) fizer uso de um filtro cromático, gerará imagens com um viés sistemático de cor. Talvez o exemplo mais crítico seja o de câmeras com baixa resolução utilizadas para fins de reconhecimento facial, pois essas podem gerar, juntamente com o viés de amostra, um viés de discriminação.

Por fim, analisa-se o viés de cognição, que ocorre, segundo alguns autores, porque o processo de tomada de decisões pode ser feito pelo uso de regras simples e práticas (heurísticas) que se caracterizam como alternativas simples no trato de grande quantidade de informações, na escassez de tempo e normalmente feito de forma inconsciente (FEITOSA, 2020; KAHNEMAN, 2011), levando a aplicações inadequadas destas regras simplificadoras, de forma que isso propicia o cometimento de uma gama de erros sistemáticos e previsíveis, ou

⁵ Na Estatística, define-se amostra como sendo um subconjunto ou um grupo menor dentro de uma população e, população, o conjunto de todos os elementos que possuem um parâmetro comum; portanto, para que uma amostra seja representativa da população com determinada característica, há que se efetuar adequadamente a escolha da amostra para que represente a população, sendo a amostragem probabilística um dos métodos utilizados para tanto. Neste método, os elementos da amostra devem ser tais que a possibilidade de qualquer deles ser selecionado seja a mesma que a possibilidade de selecionar qualquer dos elementos de uma população (TAVARES, 2007).



seja, o viés de cognição surge devido a aplicação inadequada de uma heurística numa tomada de decisão.

O viés de preconceito, uma espécie de viés de cognição, ocorre quando os dados utilizados são influenciados por preconceitos/estereótipos culturais provenientes da população, pois trata-se de uma tentativa do cérebro de simplificar o processamento de informações por meio de regras simples, atribuindo determinadas características a todas as pessoas com alguma semelhança. Há autores, porém, que defendem que vieses de cognição ocorrem porque os processos de tomadas de decisões, principalmente em temas moralmente carregados, sofrem influências exógenas, sendo, portanto, necessária uma compreensão mais aprofundada sobre o funcionamento da mente (HORTA, 2019). Esse viés pode ser considerado o principal causador do viés de algoritmo, uma vez que o viés de cognição influencia sobremaneira a elaboração do programa de IA.

Para ilustrar com algumas IA's que apresentaram vieses e que se tornaram muito comentadas devido a este fator, pode-se citar o *software* da Amazon desenvolvido para a triagem de currículos no processo de recrutamento de novos funcionários. O algoritmo criado deveria atribuir estrelas para os currículos de acordo com a qualificação do(a) candidato(a). Após o teste do software, a empresa percebeu que o algoritmo reproduzia um viés cognitivo humano atribuindo um coeficiente negativo a termos associados com mulheres (COWGILL, 2020). Pode-se mencionar, outrossim, o software de reconhecimento de imagens da Google, que classificava pessoas negras como gorilas (BBC, 2015); o *COMPAS: Correctional Offender Management Profiling for Alternative Sanctions*, um *software* desenvolvido por uma empresa privada e utilizado por algumas cortes americanas, como a do Estado de Wisconsin, para avaliar o grau de periculosidade e de possibilidade de reincidência de um réu para, assim, determinar a probabilidade de reincidência de prisioneiros. Uma empresa jornalística demonstrou, após um estudo, que o algoritmo possuía um viés racista. O *software* fornecia um escore para o cometimento de crimes como apresentando maior risco para as pessoas negras e menor para as brancas (VIEIRA, 2019).

Há diferentes propostas para a minimização ou correção dos vieses em IA. Essas propostas são muito parecidas, ou, em alguns casos, idênticas às existentes para a solução desses vieses fora do mundo da IA. Tais propostas dependem de qual ou quais vieses se pretende corrigir. A seguir, são apresentadas algumas propostas para solucionar os principais vieses apontados acima.



4 POSSÍVEIS SOLUÇÕES PARA A CORREÇÃO DOS VIESES

Os vieses gritantes, ou grosseiros, podem ser identificados inicialmente e descartados com facilidade, já no ato da programação ou na análise preliminar dos dados. Porém, aqueles mais sutis, ou menos evidentes, só podem ser observados após o uso da IA, exigindo um cuidado muito grande para que não promovam resultados equivocados e para que não percam a credibilidade mesmo após a correção.

As propostas de correção dos vieses de algoritmo aplicam-se, também, aos vieses de cognição. Uma vez que eles são inerentes ao ser humano, podem ser mitigados através da participação de uma pluralidade de agentes para atuar nas diferentes etapas do processo de criação e de execução da IA. É necessária a entrada imediata de correção para eliminar o viés assim que ele for diagnosticado.

Uma solução para os vieses de amostra também baseada na diversidade consiste em utilizar bases de dados bem diversificadas, pois, assim, é possível evitar a introdução de vieses por meio de dados não representativos ou tendenciosos. O viés de amostra é bastante crítico nas situações em que previsões ou decisões devam ser tomadas com base em correlações ou em informações indiretas, intermediadas (*proxies*). A correlação fornece informação sobre a interdependência de duas variáveis⁶. Assim, duas variáveis correlacionadas não necessariamente estão ligadas pela causalidade, destacando-se que correlação não é sinônimo de causalidade. A causalidade é caracterizada pela ligação existente entre causa e efeito; duas variáveis correlacionadas não necessariamente estão ligadas pela causalidade. Com o fim de melhor elucidar essa diferença, conta-se que o dinamarquês Gustav Fischer (PIGNOTTIA, 2011) teria provado, na década de 1930, a existência de correlação positiva entre o tamanho da população da cidade de Oldenburg e o número de cegonhas. Na realidade, Fischer só teria provado que a população e o número de cegonhas aumentaram ao longo do período de estudo. O resultado não significa que o

⁶ O valor da correlação entre duas variáveis pode variar entre -1 a 1. Quando a correlação é positiva, diz-se que há uma relação entre elas no mesmo sentido, ou seja, quando uma cresce a outra também, e quando é negativa, que quando uma cresce a outra diminui e vice-versa. Quando zero, diz-se que não há relação entre as variáveis. Uma correlação é fraca quanto mais próxima de zero for seu valor e forte, quanto mais próximo de 1(-1) for seu valor.



crescente número de cegonhas teria causado o aumento observado na população, a não ser que se acredite que são as cegonhas que trazem os bebês.

A inferência por meio de correlação constitui-se num dos fatores mais importantes de discriminação/preconceito no uso de IA (O'NEIL, 2020). A utilização da IA para a análise de grande quantidade de dados exacerba eventuais vieses já existentes, como em situações em que se usam dados como o endereço e CEP - *geopricing* e *geoblocking* - como ocorreu no caso da plataforma Decolar (OLIVA, 2021), em que a empresa cobrava mais caro de brasileiros por acomodações, assim como negava acesso a vagas, com base na localização geográfica do consumidor ou ainda, casos de precificação de prêmios de seguros de acidentes de automóveis com base na renda salarial. Nesse caso, configura-se uma discriminação indireta (OLIVA, 2021; CORBO, 2017), devida a forte correlação entre essas variáveis e o índice de criminalidade na região onde reside o segurado. Com isso, quando a correlação por meio dessas variáveis intermediárias são utilizadas, o seguro de acidentes automobilísticos é sistematicamente mais caro para moradores de regiões em que residem pessoas com menor poder aquisitivo, mesmo que o histórico do segurado não demonstre grande risco de que ele sofra ou cause acidentes de trânsito, inserindo um viés de preconceito para com aquelas que residem em regiões com menor valorização imobiliária, o que também se considera uma discriminação estatística (FRAZÃO, 2021).

A solução para a correção dos vieses, nesses casos, é a análise de variáveis que realmente sejam relevantes para a situação, evitando o uso de correlações entre variáveis intermediárias, e, mesmo quando ocorram correlações fortes, proceder à análise de ocorrência de causalidade ou não entre essas variáveis, o que somente pode ser feito por humanos, já que, como inicialmente exposto, o processo de decisão/previsão, na IA, é calcado na análise de formação de padrões.

Fora esses aspectos, o tamanho das bases de dados iniciais, que, no caso da IA, podem ser consideradas como a amostra do ponto de vista estatístico, é relevante para que ela possa abranger a maior diversidade possível, uma vez que as bases de dados devem conter elementos que descrevam significativamente, não somente em quantidade, mas também em qualidade, o universo de interesse. Na hipótese de o viés decorrer de deficiências na base de dados por falta de representatividade, sua eliminação dependerá do aumento da base ou da incorporação de dados que a tornem significativa do espectro do universo de interesse.



O viés de medida é, tecnicamente, o de mais fácil solução, haja vista que envolve somente a simples troca ou recalibração do dispositivo de medição.

Em derradeiro, aponta-se que o viés de cognição, em geral, pode ser solucionado ou mitigado por meio da ação de uma pluralidade de agentes atuando/discutindo/controlando os vários estágios envolvidos na IA. Isso porque, como já explanado alhures, o viés de cognição é inerente ao ser humano, pois toda hermenêutica (ato de interpretação) sofre a influência dos preconceitos (pré-compreensões do agente derivada de suas experiências) e da tradição (influência da sociedade, das gerações), ou seja, sempre sofre influência dos aspectos subjetivos, conforme ensina Gadamer (GADAMER, 1999). Para tentar amenizar esta influência, o autor indica como caminho o esforço do intérprete em trazer à consciência estas influências que sofre para, racionalmente, tentar reduzi-la. Porém, ele reconhece que este processo é gradativo, ascendente, pelo que o denomina “espiral da hermenêutica”. A espiral da hermenêutica está intrinsecamente ligada ao também exposto processo de formação de conceitos/soluções relacionados aos sistemas de cognições 1 e 2, também conhecidos por teorias do processo dual (FEITOSA, 2020; HORTA, 2019; KAHNEMAN, 2011; ONUMULHERES, 2016). Como ainda não é possível criar uma IA capaz de executar essa espiral da hermenêutica sem a intervenção humana, visto que o raciocínio/pensar da IA está baseada na análise de dados e a partir dessa análise, identificar padrões, comparando-os com os novos dados, a IA é incapaz de produzir como previsão/resultado algo que já não exista na sua base ou que seja composição de soluções já existentes. Portanto, até o presente, a IA não possui capacidade criativa, logo solucionar/mitigar esse viés só é possível por meio da democratização dos processos envolvendo a IA.

Neste aspecto, há que se concordar com os monistas (QUARESMA, 2019; QUARESMA, 2020; SEARLE, 1997) quando sustentam que a inexistência da consciência faz ser impossível, ao menos até o momento, se falar em uma Inteligência Artificial nos moldes do que se tem por meio de processos biológicos, haja vista, a inviabilidade de se reproduzi-la. Com isso, aufere-se a impropriedade da IA para executar a “espiral hermenêutica”, fazendo com que a metáfora baseada no antropomorfismo invalide o nome IA.

De fato, o elemento comum aos diversos vieses que podem estar presentes numa IA é sua origem por ação inevitável do ser humano, seja por vieses que o homem possua intrinsecamente ou por erros cometidos na hora da criação do algoritmo, na calibração do instrumento de captação de dados, na escolha dos dados utilizados para a alimentação da base



de dados a ser utilizada ou na interpretação das previsões/soluções, e, assim, a solução/mitigação do(s) vies(es) só pode(m) ser feita(s) por meio da ação do ser humano de forma contínua e permanente.

Fica candente, portanto, a necessidade de transparência, ampla divulgação dos algoritmos e dos processos usados para a produção de uma análise via IA, tornando necessário o controle permanente da comunidade política (cidadãos, instituições, mídia, etc.) (BAER, 2019; LEE, 2019) sobre possíveis vieses e o aprimoramento, também contínuo, dos elementos que podem ser posteriormente inseridos para fins de melhoria da qualidade da decisão, principalmente, com o auxílio de novas tecnologias.

Nos dias atuais, o termo utilizado para expressar o conceito que denota práticas a serem observadas nesse caso, é *accountability*, expressando o agir pautado por transparência das ações, responsabilidade ética, com uma devida e adequada prestação de contas de tais atos, estando ligada à ideia de governança e também à de responsabilidade civil (PINTO, 2020). Nesses termos, a IA pode ser um importante mecanismo de controle das decisões judiciais (ROQUE, 2021).

5 SUGESTÕES DE USO DA IA NO DIREITO

Diante da constatação da inevitabilidade dos vieses na IA após a revisão da literatura própria, persiste o questionamento sobre a possibilidade de emprego dela na seara jurídica, como instrumento decisório.

Quando se trabalha com uma quantidade de dados muito grande, com estudo de fenômenos das Ciências Naturais e Engenharias, entre outros, os resultados fornecidos pela IA estão menos propensos a influência de vieses que quando trabalhados pela inteligência natural, haja vista que os algoritmos, enquanto comandos a serem executados para a solução de problemas nessas áreas, já estão naturalmente adequados para o uso da IA. De fato, a linguagem com que a natureza se expressa para os humanos é a matemática; logo, o algoritmo a ser utilizado já está encaminhado - o que minimiza as possibilidades de vieses na IA.

Nas Ciências Humanas ou Sociais (Aplicadas), o problema inerente à presença de vieses na IA pode ser mais acentuado. Nesse universo, a compreensão (GADAMER, 1999) ocupa papel de relevância; logo, intrínseca a interação existente entre o “objeto” a ser analisado e o responsável pela análise. Isso faz com que se preconize o uso da IA como um



instrumento norteador de decisões, mas não seja ainda seguro atribuir natureza decisória ao resultado obtido por meio do uso dela.

Um exemplo dessa situação é o uso de IA como norteadora da rota mais célere que um motorista pode seguir, como no aplicativo *Waze*. A IA indica um caminho, mas o motorista tem a possibilidade de escolher seguir ou não a indicação, pois como a base de dados utilizada ainda não fornece informações precisas sobre a segurança do trajeto, e, em princípio, poderia indicar um caminho não seguro ao motorista ou não indicar outro seguro, mas considerado inseguro pelo aplicativo – por conta de um viés de preconceito (VELLEDA, 2019) - portanto necessário que a decisão de acatamento ou não da sugestão do aplicativo ainda deva ser tomada com cuidado pelo motorista, sem que isto elimine, contudo, a utilidade do aplicativo na apresentação de alternativas.

A IA pode, ainda, ser muito útil no mundo jurídico, diante do elevado número de ações que tramitam em juízo. Segundo amplamente noticiado, inclusive nos dados do relatório produzido pelo Conselho Nacional de Justiça, Justiça em Números 2021 (BRASIL, 2021), o poder judiciário, no ano de 2020, encerrou o ano com 75,4 milhões de processos em tramitação, dos quais 62,4 milhões estavam em andamento. No ano de 2020, segundo o mesmo relatório, o poder judiciário julgou 27,9 milhões de ações, enquanto foram registrados 25,8 milhões de novos processos.

Desses, 17,6 milhões foram ações originárias. Muitas dessas ações possuem objetos semelhantes, e, como já explanado, devido a grande quantidade de dados, a IA pode executar tarefas como classificar e catalogar de forma mais rápida e precisa que o ser humano. Por isso, pode ser um eficiente assistente do juiz, produzindo minutas de decisões (ROQUE, 2021) para a análise criteriosa dos magistrados, capaz de mitigar vieses que estejam presentes nas decisões ou previsões produzidas pelas IAs.

De modo diverso das decisões no processo administrativo, que possuem a característica da sindicabilidade, a decisão judicial apresenta caráter de definitividade - por exemplo, uma decisão de um tribunal supremo, ou ainda na coisa julgada material -, com a estabilização do processo, como afirma Daniel Assumpção Neves (NEVES, 2017, p.85):

Afirma-se que a solução do conflito por meio jurisdicional é a única que se torna definitiva e imutável, sendo considerada a derradeira e incontestável solução do caso concreto. Essa definitividade significa que a decisão que solucionou o conflito deverá ser respeitada por todos: partes, juiz do processo, Poder Judiciário e até mesmo por outros Poderes. Em razão do desenvolvimento desse raciocínio, a coisa julgada material é fenômeno privativo das decisões jurisdicionais.



Isto torna a análise por parte dos magistrados imperiosa, pois uma decisão (sentença ou acórdão) exarada por IA com algum viés poderia gerar uma prestação deficiente por parte do Estado, haja vista a possibilidade de configurar um erro judiciário, podendo inclusive ensejar uma responsabilização civil do Estado por erro judiciário (DE ALMEIDA, 2012).

Há outro ponto que impede a prolação das decisões judiciais pela IA. Uma vez que, como explanado, nos sistemas que utilizam IAs que fazem uso de *Deep Learning* –utilizada na maioria dos sistemas preditivos ou decisórios – são opacos, então se a IA proferisse uma decisão de mérito, ela deveria, em princípio, indicar os fundamentos que justifiquem seu convencimento formado através da análise das provas produzidas no processo - como deve fazer um juiz -, construindo em contraditório seu conhecimento a respeito dos fatos da causa (a valoração democrática da prova), uma fundamentação que demonstre, discursivamente, como chegou às suas conclusões acerca da apreciação da prova, a fim de se demonstrar que a decisão proferida é a decisão correta para o caso concreto em exame, sem que isso resulte de discricionariedade ou voluntarismo judicial - teoria hermenêutica. Porém se não se é possível traçar uma linha lógica de como o *input* (provas e evidências) levam ao *output* (decisão ou previsão), há a ausência da teoria hermenêutica escolhida que dê suporte a metodologia interpretativa adotada e isso, viola o disposto no art. 371 do CPC/2015, que dispõe: “o juiz apreciará a prova constante dos autos independentemente do sujeito que a tiver promovido, e indicará na decisão as razões da formação de seu convencimento”, conhecido como princípio do livre convencimento motivado ou persuasão racional, adotado pela jurisprudência pátria (BRASIL, 2022).

A crítica aos vieses que possam estar presentes nas IA's exige muita atenção quando de seu emprego na seara jurisdicional, mas não se pode esquecer que os magistrados, enquanto humanos, também apresentam vieses, e que, nas decisões das causas rotineiras, podem agir mediante heurística, pois o sistema cognitivo 1 também passa a atuar de forma mais intensa quando há situações que, de tão frequentes e similares, induzem a ações que deixem de lado o sistema cognitivo 2. Assim sendo, o uso da IA pode, até mesmo, propiciar o diagnóstico de vieses (SUNSTEIN, 2018) nas decisões judiciais e auxiliar na correção ou minimização deles.

Os jurisdicionados podem se beneficiar sobremaneira do emprego da IA nas ações em que se discute o direito à saúde, tendo em vista o grande volume de ações impetradas anualmente. Segundo o Relatório Analítico Propositivo Justiça Pesquisa: Judicialização da



Saúde no Brasil: Perfil das demandas, causas e propostas de solução, 2019 (BRASIL, 2019), entre 2008 a 2017, constataram-se, aproximadamente, 770 mil demandas desse jaez. A categoria “fornecimento de medicamentos”, mesmo que associado a outros termos, é predominante.

Segundo Duarte e Pimenta (DUARTE, 2020), entre várias razões que levam à judicialização da saúde, as políticas públicas deficitárias são uma das principais. No Brasil, em virtude das recentes alterações normativas que implicam constrição nas políticas públicas sanitárias, não há perspectiva de diminuição nos números da judicialização da saúde. Dessa forma, a IA pode ser de grande valia pela sua velocidade de análise e precisão, podendo diminuir o tempo de prolação de sentenças em primeira instância ou de acórdãos em segunda - ponto crucial em casos de urgência e/ou emergência - por meio da proposição de minutas de decisões, inclusive mitigando ou até resolvendo um grande problema que ocorre nas ações de judicialização da saúde, qual seja, a ausência de motivação específica e decisões menos díspares para situações semelhantes e em consonância com precedentes, pois de acordo com Didier (2016, p. 484),

É preciso - e exigível - que a decisão judicial identifique exatamente as questões de fato que se reputaram como essenciais ao deslinde da causa e delineie, também de forma explícita, a tese jurídica adotada para a sua análise e para se chegar à conclusão exposta na parte dispositiva. É também preciso - e, igualmente, exigível - que, ao aplicar ou deixar de aplicar um precedente, o órgão jurisdicional avalie, de modo explícito, a pertinência da sua aplicação, ou não, ao caso concreto, contrapondo as circunstâncias de fato envolvidas aqui e ali e verifique se a tese jurídica adotada outrora é adequada, ou não, para o caso em julgamento. Tudo isso compõe expressamente o dever judicial de fundamentação na forma dos arts. 489, §1º, e 927, §1º, CPC.

A IA tem sido utilizado em cortes brasileiras (TOLEDO, 2021; ROQUE, 2021) ao longo dos últimos anos, e, verifica-se um crescente interesse e discussões sobre o uso da IA no Direito e suas repercussões, como pode ser observado nas publicações resultados de congressos como os CONPEDI dos anos de 2020 e 2021 (CONPEDI, 2020; CONPEDI, 2021).

Como perspectiva futura, vislumbra-se sua possibilidade de emprego na judicialização da saúde, tanto com o objetivo de diagnosticar as colisões entre direitos fundamentais presentes nestes processos judiciais, conforme proposto por Alexy (ALEXY, 2015) e explicitada para alguns casos por Duarte (DUARTE, 2020), através da proposição de possíveis leis de colisão.



Reside, na colaboração entre o homem e a máquina, uma grande possibilidade de interação da IA com os juízes humanos, na parceria denominada de corrobótica (*co-robotics*) - computação cognitiva (*cognitive computing*), em que o trabalho é efetuado de forma colaborativa entre máquinas e humanos, com cada um realizando o tipo de atividade que melhor é capaz de desempenhar (BOEING, 2019).

6 CONCLUSÕES

A IA é uma ferramenta desenvolvida graças ao avanço da Ciência, não somente da Ciência dura, mas em conjunto com as Ciências Humanas, da Vida e Sociais, pois para se atingir o ponto em que se encontra hoje, foi necessária a estreita colaboração multidisciplinar e longínqua pesquisa envolvendo desde o entendimento dos processos de aprendizagem, mecânica do processamento de informações no cérebro à programação de algoritmos, desenvolvimento de novas tecnologias que suportem um grande processamento de dados, entre outras.

Sendo desenvolvida por seres humanos, mesmo que inconscientemente, também incorpora heurísticas e vieses característicos dos homens, acrescido da circunstância agravante caracterizada pela inviabilidade de que a IA execute a espiral da hermenêutica de forma a mitigar esses vícios. Necessário, portanto, um programa de *accountability* que permita a governança das ações da IA.

Em virtude, contudo, da inevitabilidade dos vieses na IA, para que eles não produzam implicações aos jurisdicionados antes da detecção e correção dos vieses, considera-se que, no âmbito do exercício das competências jurisdicionais, a IA não deve ser utilizada na prolação final de sentenças, pois a falta da transparência, frente a ausência da teoria hermenêutica escolhida pelo julgador que dê suporte a metodologia interpretativa adotada, violaria o princípio da motivação e do livre convencimento motivado ou persuasão racional.

Contudo, a IA pode ser utilizada, judicialmente, para minutar uma decisão em auxílio ao magistrado, uma vez que estaria sujeita a cometer menos erros na análise do grande volume de processos que um ser humano pode ensejar por motivos como cansaço e heurísticas devido à semelhança dos processos. A IA e o ser humano podem trabalhar em conjunto – corrobótica - de maneira a incrementar a eficiência da atuação jurisdicional,



sobretudo para a análise de grande volume de processos repetitivos, como constatado na judicialização da saúde.

A IA ainda pode ser útil para a descoberta de vieses dos magistrados, uma vez que poderia amplificá-lo e, assim, torná-lo mais evidente e passível de correção (SUNSTEIN, 2018). Portanto, a existência de vieses na IA não impede que ela seja utilizada em benefício do aumento da qualidade de vida humana sequer na arena jurisdicional.

L.G.M.D. agradece o apoio financeiro recebido do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

7 REFERÊNCIAS

ALEXY, Robert. *Teoria dos direitos fundamentais*. Tradução de Virgílio Afonso da Silva. 2. Edição. São Paulo. Malheiros, 2015.

BAER, Tobias. *Understand, Manage, and Prevent Algorithmic Bias: A Guide for Business Users and Data Scientists*. Berkeley, Apress, 2019.

BBC. *Google apologises for Photos app's racist blunder*. 2015. Disponível em: <https://www.bbc.com/news/technology-33347866>. Acesso em: 30 de ago. 2021.

BIBLIOTECONOMIADIGITAL. *BASES DE DADOS: conceito, classificações, critérios, aspectos importantes e exemplos*. Biblioteconomiadigital, 2017. Disponível em: <https://biblioteconomiadigital.com.br/2017/12/bases-de-dados-conceito-classificacoes.html#:~:text=Base%20de%20dados%3A%20conjunto%20de,conjunto%20de%20%20Bases%20de%20Dados>. Acesso em: 07 out. 2021.

BOEING, D. H. A. B. *Ensinando um robô a julgar: pragmática, discricionabilidade e vieses no uso de aprendizado de máquina no judiciário*. 2019. Trabalho de Conclusão de Curso de Graduação em Direito do Centro de Ciências Jurídicas da Universidade Federal de Santa Catarina, Florianópolis, 2019.

BRASIL. Conselho Nacional de Justiça. *Relatório Analítico propositivo Justiça Pesquisa: Judicialização da Saúde no Brasil: Perfil das demandas, causas e propostas de solução*. CNJ, 2019. Disponível em: <https://www.cnj.jus.br/wp-content/uploads/2011/02/95da70941b7cd226f9835d56017d08f4.pdf>. Acesso em: 18 nov. 2021.



_____. Conselho Nacional de Justiça. *Justiça em números 2021*, Brasília: CNJ, 2021.

Disponível em:

<https://www.cnj.jus.br/wp-content/uploads/2021/09/relatorio-justica-em-numeros2021-12.pdf>.

Acesso em 18 nov. 2021.

_____. Supremo Tribunal Federal - ARE: 1335473 MG 1029171-53.2019.4.01.0000,

Relator: EDSON FACHIN, Data de Julgamento: 11/11/2021, Segunda Turma, Data de

Publicação: 10/01/2022.

CANZIANI, Luísa. *Marco Legal da Inteligência Artificial - as três premissas*. Correio Braziliense, 2021. Disponível em:

<https://www.correiobraziliense.com.br/opiniaio/2021/08/4942488-artigo-marco-legal-da-inteligencia-artificial---as-tres-premissas.html>. Acesso em: 10 ago. 2021.

CARVALHO, A.C.P.L.F. *Redes Neurais Artificiais*. ICMC. [s.d]. Disponível em:

<https://sites.icmc.usp.br/andre/research/neural/>. Acesso em: 24 abr 2022.

CONPEDI. *Congresso Internacional: Direito e Inteligência Artificial*. Conpedi. 2020.

Disponível em:

<https://www.conpedi.org.br/wp-content/uploads/2020/09/SKEMA-Intelig%C3%Aancia-Artificial-e-tecnologias-aplicadas-ao-Direito-IV.pdf>. Acesso em: 24 abr 2022.

_____. II CONGRESSO INTERNACIONAL DE DIREITO E INTELIGÊNCIA ARTIFICIAL. Conpedi. 2021. Disponível em:

<http://site.conpedi.org.br/publicacoes/b3vv7r7g>. Acesso em: 24 abr 2022.

CORBO, W. *Discriminação Indireta: o que é e como superá-la?*. Jota, 2017. Disponível em:

<https://www.jota.info/opiniaio-e-analise/artigos/discriminacao-indireta-o-que-e-e-como-supera-la-09112017>. Acesso em: 15 out. 2021.

COWGILL, Bo; TUCKER, Catherine E. *Algorithmic Fairness and Economics*. Columbia Business School Research Paper. 2020. Disponível em:

<https://dx.doi.org/10.2139/ssrn.3361280>. Acesso em: 30 de ago. 2021.

DE ALMEIDA, Vitor Luís. *A responsabilidade civil do Estado por erro judiciário sob a ótica do sistema lusófono. Análise nos ordenamentos jurídicos português e brasileiro*. Revista de Informação Legislativa, v. 49, n196. 2012. Disponível em:

https://www12.senado.leg.br/ril/edicoes/49/196/ril_v49_n196_p259.pdf. Acesso em: 25 abr. 2022.



DIDIER JUNIOR, Fredie; BRAGA, Paula Sarno; OLIVEIRA, Rafael. *Curso de direito processual civil: Teoria Geral da Prova, Direito Probatório, Decisão, Precedente, Coisa Julgada e Tutela Provisória*, v. 2, 11ª ed. Salvador. Ed. JusPodivm, 2016.

DUARTE, Luciana G. M.; PIMENTA, Liana de B., A tutela jurisdicional do direito à saúde: Uma análise do relatório analítico propositivo “Judicialização da saúde no Brasil: Perfil das demandas, causas e propostas de solução”, do Conselho Nacional de Justiça. In *Direito à saúde: Judicialização e pandemia do novo coronavírus*, coordenação: L. G. M. Duarte e V. L. Vidal. São Paulo. Ed. Thomson Reuters, 2020.

DUARTE, Luciana. G. M. *POSSIBILIDADES E LIMITES DO CONTROLE JUDICIAL SOBRE AS POLÍTICAS PÚBLICAS DE SAÚDE: Um Contributo para a Dogmática do Direito à Saúde*. 2. Edição. Belo Horizonte. Editora Forum, 2020.

FEITOSA, Arnaldo B. ; ROGERS, Pablo; ROGERS, Dany. *A influência de vieses cognitivos e motivacionais na tomada de decisão gerencial: evidências empíricas em uma empresa de construção civil brasileira*. Revista de Negócios, Blumenau, v. 19, n. 3, p. 3-22, jul/set, 2020.

FRAZÃO, Ana. *Geopricing e geoblocking: as novas formas de discriminação de consumidores*. Jota, 2018. Disponível em:

<https://www.jota.info/opiniao-e-analise/colunas/constituicao-empresa-e-mercado/geopricing-e-geoblocking-as-novas-formas-de-discriminacao-de-consumidores-15082018>. Acesso em 14 out. 2021.

_____. *Discriminação algorítmica: por que algoritmos preocupam quando acertam e erram?*. Jota, 2021. Disponível em:

<https://www.jota.info/opiniao-e-analise/colunas/constituicao-empresa-e-mercado/discriminacao-algoritmica-por-que-algoritmos-preocupam-quando-acertam-e-erram-04082021>. Acesso em: 14 out. 2021.

GADAMER, Hans-Georg, *Verdade e Método: Traços fundamentais de uma hermenêutica filosófica*. Tradução Flávio Paulo Meurer, Petrópolis, 3ª. ed, Editora Vozes, 1999.

GOMES, Pedro C. T. *Ética e Inteligência Artificial: viés em machine learning*. Datageeks, 2019. Disponível em: <https://www.datageeks.com.br/etica-e-inteligencia-artificial/>. Acesso em: 06 de ago. 2021.



HORTA, Ricardo de L. *Por que existem vieses cognitivos na Tomada de Decisão Judicial? A contribuição da Psicologia e das Neurociências para o debate jurídico*. Revista Brasileira de Políticas Públicas, Brasília, v.9, n. 3, p. 83-122, 2019.

KAHNEMAN, D. *Rápido e Devagar: Duas formas de Pensar*, tradução de Cássio de Arantes Leite. Rio de Janeiro, Objetiva, 2011.

KANDEL, E.R.; KOESTER, J. D.; MACK, S. and SIEGELBAUM, S. *Principles of neural science*. McGraw-hill New York, 6 edition, 2021.

LEE, N. T. ; RESNICK, P. ; BARTON, G. *Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms*. Brookings, 2019. Disponível em: <https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/>. Acesso em: 22 set. 2021.

LOURENCINI, Antônio Rogério; DA COSTA, Yvete Flávio. *O conteúdo mínimo da fundamentação das decisões judiciais no ordenamento jurídico brasileiro*. Revista da Faculdade de Direito UFPR, v. 63, n. 1, p. 161-187, 2018.

MANYIKA, James; SILBERG, Jake. *Como lidar com vieses na inteligência artificial (e nos seres humanos)*. McKinsey Global Institute, 2019. <https://www.mckinsey.com/featured-insights/artificial-intelligence/tackling-bias-in-artificial-intelligence-and-in-humans/pt-br>. Acesso em: 07 ago. 2021.

MICHAELIS. *Bias*. Michaelis. Disponível em: <https://michaelis.uol.com.br/moderno-portugues/busca/portugues-brasileiro/bias>. Acesso em: 25 ago. 2021.

MORETTIN, Pedro A.; BUSSAB, Wilton de O. *Estatística básica*. São Paulo, Saraiva, 2010.

NEGRI, Sérgio M. C. A. *Robôs como pessoas: a personalidade eletrônica na Robótica e na inteligência artificial*. Pensar, Revista de Ciências Jurídicas, Fortaleza, v. 25, p. 1-14, jul/set, 2020.

NUNES, Dierle; MARQUES, Ana L. P. C. *Inteligência artificial e direito processual: vieses algorítmicos e os riscos de atribuição de função decisória às máquinas*. Revista de Processo, v. 285, p. 421-447, nov, 2018.

OLIVA, Milena D.; SILVA, Jenifer G. da. *Discriminação algorítmica nas relações de consumo*. Migalhas, 2021. Disponível em:



<https://www.migalhas.com.br/depeso/340680/discriminacao-algoritmica-nas-relacoes-de-cons-umo>. Acesso em: 14 out. 2021.

O'NEIL, Cathy. *Algoritmos de destruição em massa: como o big data aumenta a desigualdade e ameaça a democracia*. Tradução Rafael Abraham. Santo André, 1^a. ed, Ed. Rua do Sabão, 2020.

ONUMULHERES. *Vieses inconscientes, equidade de gênero e o mundo corporativo: lições da oficina “vieses inconscientes”*. Onumulheres, 2016. Disponível em:

https://www.onumulheres.org.br/wp-content/uploads/2016/04/Vieses_inconscientes_16_digital.pdf. Acesso em: 30 de ago. 2021.

ORACLE. *O que é inteligência artificial - IA?*. Oracle. Disponível em:

<https://www.oracle.com/br/artificial-intelligence/what-is-ai/>, Acesso em: 07 ago. 2021.

PARLAMENTO EUROPEU. *Resolução do Parlamento Europeu, de 16 de fevereiro de 2017, que contém recomendações à Comissão sobre disposições de Direito Civil sobre Robótica*.

Bruxelas, 2017. Disponível em:

https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_PT.pdf. Acesso em: 19 out. 2021.

PIGNOTTIA, Angela. *Stork Population vs. Births*. Pignottia, 2011. Disponível em:

<http://pignottia.faculty.mjc.edu/math134/classnotes/storks.pdf>. Acesso em: 19 set. 2021.

PRADO, Jean. *A Inteligência artificial é mais antiga do que você imagina*. Tecnoblog, 2016.

Disponível em: <https://tecnoblog.net/195106/inteligencia-artificial-historia-dilemas/>. Acesso em: 09 ago. 2021.

PINTO, Henrique Alves. *A utilização da inteligência artificial no processo de tomada de decisões: por uma necessária accountability*. Revista de Informação Legislativa, v. 57, n. 225, p. 43-60, 2020.

PROGRA{M}ARIA. *Quais são os tipos de aplicações de Inteligência Artificial mais comuns?*. Progra{m}aria, 2020. Disponível em:

<https://www.programaria.org/quais-sao-os-tipos-de-aplicacoes-de-inteligencia-artificial-mais-comuns/>. Acesso em: 18 nov. 2021.

_____. *Algoritmos de Inteligência Artificial (IA) e Vieses: uma reflexão sobre ética e justiça*.

Progra{m}aria, 2020. Disponível em:

<https://www.programaria.org/algoritmos-de-inteligencia-artificial-e-vieses-uma-reflexao-sobre-etica-e-justica/>. Acesso em: 10 ago. 2021.



- QUARESMA, Alexandre. *Inteligências artificiais e o problema da consciência*, Paakat: Revista de Tecnología y Sociedad, Universidad de Guadalajara. 2019. Disponível em: <http://dx.doi.org/10.32870/Pk.a9n16.349>. Acesso em 09 ago. 2021.
- _____. Intencionalidade e Inteligência Artificial, *Paakat: Revista de Tecnología y Sociedad*, Universidad de Guadalajara. 2020. Disponível em: DOI: <http://dx.doi.org/10.32870/Pk.a10n18.403>. Acesso em 09 ago. 2021.
- ROQUE, Andre; DOS SANTOS, Lucas Braz Rodrigues. *Inteligência artificial na tomada de decisões judiciais: três premissas básicas*. Revista Eletrônica de Direito Processual, v. 22, n. 1, 2021.
- SEARLE, John R. *The mystery of consciousness*, New York, The New York Review of Books, 1997.
- SUNSTEIN, Cass R. *Algorithms, Correcting Biases*. Preliminary draft for Social Research, 12.1.2018. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3300171. Acesso em: 06 ago. 2021.
- TAVARES, Marcelo. *Estatística Aplicada à Administração*. Florianópolis, Ed. UFSC, 2007.
- TOLEDO, Cláudia; ALVES, Michelle. *Inteligência Artificial Aplicada ao Direito*. In: I Seminário do Observatório de Práticas Sociojurídicas, jun. 2021. UFERSA. Disponível em: <https://ppgd.ufersa.edu.br/wp-content/uploads/sites/168/2021/05/Lista-de-Trabalhos-Aprovados-nos-GTs-do-I-Seminario-de-Pesquisa-do-OPS.pdf>. Acesso em: 10 mar. 2022.
- VELLEDA, Luciano. *'Alerta de risco de crime' do Waze provoca debate sobre preconceito em São Paulo*. RBA. 2019. Disponível em: <https://www.redebrasilatual.com.br/cidadania/2019/03/alerta-de-risco-de-crime-do-waze-provoca-debate-sobre-preconceito-em-sao-paulo/>. Acesso em: 24/04/2022
- VIEIRA, Leonardo M. *A problemática da inteligência artificial e dos vieses algorítmicos: caso COMPAS*. In: 2019 Brazilian Technology Symposium. Unicamp. Disponível em: <https://www.lcv.fee.unicamp.br/images/BTSym-19/Papers/090.pdf>. Acesso em: 30 ago. 2021.
- WIKIPEDIA. *História da inteligência artificial*. Wikipedia. Disponível em: https://pt.wikipedia.org/wiki/Hist%C3%B3ria_da_intelig%C3%Aancia_artificial. Acesso em: 06 ago. 2021.