DISCRIMINAÇÃO ALGORÍTMICA NA RELAÇÃO DE EMPREGO: EFICIÊNCIA ECONÔMICA, INTELIGÊNCIA ARTIFICIAL E FRAGILIDADE DO EMPREGADO

ALGORITHMIC DISCRIMINATION IN THE EMPLOYMENT RELATIONSHIP: ECONOMIC EFFICIENCY, ARTIFICIAL INTELLIGENCE AND EMPLOYEE FRAGILITY

João Luís Nogueira Matias[1]
Ricardo Antônio Maia de Morais Júnior[2]

RESUMO: A pesquisa busca verificar quais as peculiaridades da relação de emprego que a tornariam mais suscetível ou frágil quanto à ocorrência de discriminação algorítmica. Como resultados da pesquisa, verificou-se que o uso da IA depende do volume, da velocidade e do valor dos dados fornecidos pelo *Big Data*, gerando resultados tão bons quanto forem os dados. Além disso, entendeu-se que, apesar de buscar objetividade, os algoritmos podem causar discriminação por meio de um erro na programação, generalização, uso de informações sensíveis e limitação de direitos. Por fim, considerou-se que as relações de emprego possuem peculiaridades que tornam o empregado mais suscetível à discriminação algorítmica, por exemplo a desigualdade entre as partes; a fragilidade do empregado, diante de um legítimo poder de controle; e a nudez tecnológica, tendo em vista a quantidade de dados, inclusive desnecessários, em posse do empregador. Como metodologia, utilizou-se a pesquisa de natureza descritiva, exploratória e interpretativa, de cunho qualitativo, mediante análise documental e bibliográfica, valendo-se da utilização do método hipotético-dedutivo.

PALAVRAS-CHAVE: inteligência artificial; eficiência econômica; discriminação algorítmica.

ABSTRACT: The research seeks to verify which peculiarities of the employment relationship would make it more susceptible or fragile to the occurrence of algorithmic discrimination. As a result of the research, it was found that the use of AI depends on the volume, speed and value of the data provided by Big Data, generating results that are as good as the data. Furthermore, it was understood that, despite seeking objectivity, algorithms can cause discrimination through a programming error,

---

1 *Doutor em Direito Comercial pela Universidade de São Paulo; doutor em Direito Público pela Universidade Federal de Pernambuco; mestre em Direito e Desenvolvimento pela Universidade Federal do Ceará; MBA em gestão de empresas FGV/MARPE; professor titular da Universidade Federal do Ceará e do Centro Universitário 7 de Setembro – UNI7. Lattes: http://lattes.cnpq.br/8192937377585273. Orcid: https://orcid.org/0000-0002-3873-702X. E-mail: joaoluisnm@uol.com.br.*

2 *Mestre em Ordem Jurídica Constitucional pela Universidade Federal do Ceará; professor do Centro Universitário 7 de Setembro – Uni7; especialista em Direito Empresarial pela Fundação Getulio Vargas – FGV. Lattes: http://lattes.cnpq.br/3339871443217078. Orcid: https://orcid.org/0000-0001-8877-4016. E-mail: ricardomaia.moraisjunior@gmail.com.*

Recebido em: 3/4/2024
Aprovado em: 14/5/2024

generalization, use of sensitive information and limitation of rights. Finally, it was considered that employment relationships have peculiarities that make the employee more susceptible to algorithmic discrimination, for example the inequality between the parties; the fragility of the employee, in the face of a legitimate power of control; and technological nakedness, in view of the amount of data, including unnecessary ones, held by the employer. The methodology used was descriptive, exploratory and interpretative, of a qualitative nature, through documentary and bibliographic analysis, using the hypothetical-deductive method.

KEYWORDS: artificial intelligence; economic efficiency; algorithmic discrimination.

SUMÁRIO: 1 – Introdução; 2 – Inteligência Artificial, *Machine Learning* e *Big Data*; 2.1 – Esclarecimentos sobre Inteligência Artificial e *Machine Learning*; 2.2 – *Big Data*: quando aos dados são intensificados o volume, a velocidade e o valor; 3 – Discriminação algorítmica: conceito, tipologia e desafios jurídicos; 3.1 – Discriminação algorítmica e sua tipologia; 3.2 – Implicações jurídicas da discriminação algorítmica; 4 – Inteligência Artificial na relação de emprego: eficiência econômica x fragilidade e nudez tecnológica do empregado; 4.1 – Eficiência e produtividade como objetivos da Inteligência Artificial no ambiente de trabalho; 4.2 – Fragilidade e nudez tecnológica do empregado; 5 – Considerações finais; Referências.

1 – Introdução

As relações de trabalho, por serem aquelas nas quais um indivíduo (contratado) executa uma atividade em troca de remuneração, a qual é paga por outrem (seu contratante), que busca obter vantagem econômica (direta ou indireta), possuem íntima ligação com novas tecnologias. Afinal, se o serviço prestado pelo contratado não é o objetivo do contratante, mas sim o resultado econômico gerado por esse serviço, então a lógica do contratante será a de obter benefício mediante o serviço menos custoso.

Foi essa lógica que permeou a análise de Adam Smith sobre a manufatura de alfinetes, pois identificou que a divisão do trabalho (que também é um tipo de tecnologia, não digital, mas organizacional) permitia a produção em muito maior escala do que se houvesse a fabricação individual do produto (2016, p. 8-9).

Na busca pelo lucro, o contratante (ou empregador, no caso da relação de emprego) é incessantemente instado a buscar novas formas de otimizar sua produção, de reduzir seus custos e de aprimorar as atividades desempenhadas por seus contratados (empregados, no caso da relação de emprego). Exatamente essa busca pela eficiência econômica que proporciona a utilização de mecanismos de análise de grandes volumes de dados para gestão de pessoas, inclusive por meio de algoritmos de Inteligência Artificial.

Recentemente, foi noticiado que a Teleperformance, uma grande empresa de teleatendimento sediada na Colômbia, que fornece serviços para grandes empresas de tecnologia (como Apple, Amazon e Uber), teria pressionado seus empregados, em regime de *home office*, a assinarem uma declaração concordando com a instalação de câmeras em seus computadores. Essas câmeras estariam conectadas a um mecanismo de Inteligência Artificial que identificaria possíveis

violações às regras internas de troca de turno, automaticamente enviando uma fotografia dessa violação para a empresa apurar o ocorrido[3].

Outro caso, noticiado em 2019, foi da tecnologia baseada em Inteligência Artificial, desenvolvida pela empresa IBM, que permitiria monitorar os humores dos empregados, identificando quando eles estariam mais propensos a pedir demissão. O objetivo declarado dessa IA seria permitir ao empregador tomar decisões que pudessem realocar a pessoa para outra atividade ou de auxiliá-la em caso de algum problema que esteja lhe acometendo, de maneira a reter seus empregados[4].

Para além de questões como violação da privacidade e da proteção de dados dos empregados, assunto que demandaria maiores discussões em futuras pesquisas, ambas as notícias evidenciam que empregadores utilizam mecanismos de IA para monitorar seus empregados e, assim, proporcionar maior eficiência de suas atividades. Esse tipo de utilização, por exemplo ao monitorar possíveis falhas de conduta ou emoções de empregados, pode gerar situações discriminatórias, principalmente quando os critérios utilizados pelas tecnologias não são divulgados e monitorados por um terceiro.

Mecanismos baseados em IA já são amplamente utilizados para a melhoria de processos gerenciais e para atingir resultados de maneira mais eficiente, mas também em outros cenários, por exemplo no Poder Judiciário, quando o algoritmo Victor faz a triagem de ações apresentadas ao Supremo Tribunal Federal (STF) para identificar se há enquadramento em alguma tese de Repercussão Geral. Essa IA conseguiu reduzir o tempo de análise de um caso de 40 (quarenta) minutos (se fosse feita por um humano) para 5 (cinco) segundos[5].

Situações discriminatórias podem ocorrer em qualquer área do Direito, mas nas relações de emprego sua ocorrência pode ser, aparentemente, mais gravosa, o que suscita o questionamento que se faz na presente pesquisa: quais elementos próprios das relações de emprego poderiam gerar maior fragilidade do empregado quanto à ocorrência de discriminação causada por algoritmos de Inteligência Artificial?

Visando a responder a esse questionamento, pretende-se utilizar como metodologia a pesquisa de natureza descritiva, exploratória e interpretativa, de cunho qualitativo, mediante análise documental e bibliográfica, valendo-se da utilização do método hipotético-dedutivo. Inicialmente mediante a análise

3 Disponível em: https://www.nbcnews.com/tech/tech-news/big-tech-call-center-workers-face-pressure-accept-home-surveillance-n1276227. Acesso em: 3 abr. 2024.

4 Disponível em: https://www.cnbc.com/2019/04/03/ibm-ai-can-predict-with-95-percent-accuracy-which-employees-will-quit.html. Acesso em: 3 abr. 2024.

5 Disponível em: https://www.cnj.jus.br/wp-content/uploads/2022/04/inteligencia-artificial-e-a-aplicabilidade-pratica-web-2022-03-11.pdf. Acesso em: 3 abr. 2024.

dos principais tipos de mecanismos baseados em Inteligência Artificial e seus componentes, assim como sua relevância nas relações de emprego e na gestão de recursos humanos.

Em um segundo momento, buscar-se-á identificar de que maneiras podem ocorrer situações discriminatórias causadas por algoritmos de Inteligência Artificial, além de verificar seus impactos jurídicos decorrentes.

Por fim, analisar-se-ão quais os elementos específicos da relação de emprego que poderiam gerar algum tipo de fragilidade maior ao empregado quanto à ocorrência de discriminação algorítmica.

2 – Inteligência Artificial, *Machine Learning* e *Big Data*

A utilização de mecanismos de IA em diversas funcionalidades e cada vez mais acessíveis a todos proporciona muitas vantagens, pois permite que serviços ofertados por empresas ou pelo Poder Público sejam mais eficientes, por exemplo quando uma plataforma digital de *streaming* de vídeos possa sugerir conteúdos que sejam mais adequados ao perfil de cada um dos usuários.

No entanto, essa ampla utilização gera algumas impropriedades, principalmente conceituais, quando se aborda o assunto em termos não científicos. Buscando direcionar as bases teóricas da pesquisa, a seguir serão esclarecidos alguns conceitos relevantes para as demais seções do presente artigo.

2.1 – Esclarecimentos sobre Inteligência Artificial e *Machine Learning*

Em um contexto em que muito se fala, cotidianamente, sobre Inteligência Artificial, muitas vezes de maneiras equivocadas, faz-se necessário esclarecer, de certa forma, o que poderia ser considerado um mecanismo de inteligência artificial e por que seria mais adequado, ao invés disso, denominar inteligências artificiais (no plural).

Desde Allan Turing e John Haugeland (1950) já se percebia que o comportamento inteligente de um computador se dá quando este adquire o mesmo nível de *performance* humana, quando resolvendo testes cognitivos. Assim, para Turing, seria entendido como inteligente aquele computador que conseguisse, resolvendo esses testes, enganar um interrogador, ao ponto de este não conseguir diferenciar o homem da máquina.

Ao longo da segunda metade do século XX e até hoje, houve diversos avanços tecnológicos, que permitiram a experimentação da teoria apresentada por Turing em aplicações práticas, tanto que a maioria da população maneja, com certa facilidade, diversos mecanismos de IA no seu dia a dia, seja quando

utiliza os *smartphones*, seja quando navega na *internet*, dentre diversas outras situações.

No entanto, conceitualmente houve poucas modificações sobre o que se pode entender como Inteligência Artificial (apesar de que, na prática, diversas outras tecnologias foram desenvolvidas ou aprimoradas, mas todas integrando a mesma concepção teórica).

Kaplan e Haenlein definem esse termo de uma maneira simples, porém com grande abrangência, facilitando a compreensão do termo para os propósitos desta pesquisa. Para eles, Inteligência Artificial pode ser considerada a habilidade de um sistema de interpretar dados externos de maneira correta, aprender com esses dados e usar esse aprendizado para atingir objetivos específicos, mediante adaptação flexível (2018, p. 17).

Trata-se, portanto, de uma definição bem ampla, que pode abranger diversos tipos de tecnologia. Existem diversas metodologias que são consideradas IA, de acordo com diversos critérios, sendo por isso impossível generalizar tudo o que acontece com uma tecnologia de Inteligência Artificial como possível de acontecer com outra tecnologia, de metodologia diversa.

Em regra, a Inteligência Artificial é desenvolvida mediante a aplicação de algoritmos, os quais são sequências de instruções que informam ao computador o que ele deve fazer, mas não qualquer sequência, apenas aquelas que sejam suficientemente precisas para serem aceitas por um computador (Domingos, 2017, p. 24-26).

Esses algoritmos são alimentados por dados (*input*), a partir dos quais, utilizando um código matemático previamente definido, produzirão novos dados (*output*) como resultado. Simplificadamente, esse é o funcionamento do tipo mais básico daquilo que se chama hoje de Inteligência Artificial.

Essa lógica se acentua (e se modifica) quando se trata de algoritmos de aprendizado (denominados *learners*), utilizados no aprendizado de máquina (em inglês, *machine learning*).

Por meio do *machine learning*, novos algoritmos são criados pela própria IA, por meio da análise conjunta dos dados com os resultados pretendidos. Ou seja, como explica Pedro Domingos (2017, p. 29), a lógica dos algoritmos tradicionais se inverte: ao invés de receberem (*input*) dados, para que o algoritmo produza mais dados como resultado (*output*), os algoritmos *learners* recebem (*input*) os dados e o resultado pretendido, de maneira que produzam (*output*) novos algoritmos, os quais capazes de transformar um (dados) no outro (resultado).

2.2 – *Big Data*: quando aos dados são intensificados o volume, a velocidade e o valor

Partindo do entendimento de que a IA e o *machine learning* atuam baseados em dados, evidencia-se a íntima relação dessas áreas com o *Big Data*. Esse termo é frequentemente relacionado a três aspectos, volume, velocidade e valor. Isso porque o *Big Data* representa não apenas um enorme banco de dados à disposição (volume), mas também representa a velocidade de processamento desses dados, praticamente em tempo real (velocidade), bem como o valor que esses dados possuem hoje, sendo considerados como ativos econômicos para diversos modelos de negócios (valor) (Doneda, 2018, p. 5).

Reforçando essa ideia, Mayer-Schönberger e Cukier em obra seminal sobre o tema afirmam não haver uma definição precisa sobre o termo *Big Data*, mas, apesar disso, seria possível caracterizá-lo com base em três tendências. Primeiro, trata-se de uma imensa quantidade de dados, tantos que se aproximam da totalidade, pois a conexão dessas informações pode ser inimaginável. Segundo, devido à enorme quantidade de dados, eles podem estar desorganizados e, portanto, podem gerar certa imprecisão. Em terceiro lugar, com o *Big Data* é possível converter-se a causalidade em correlação (2014, p. 13-14), gerando possibilidades de utilização antes nunca vistas.

Ou seja, a clássica concepção de que o conhecimento científico deveria advir de causalidades, ou seja, buscar situações que causariam determinados efeitos, parece superada pela tomada de decisão baseada em estatísticas, ou seja, ocorrendo um evento "x" é muito provável (estatisticamente falando) que ocorra o evento "y" (Mendes; Mattiuzzo, 2019, p. 44).

Diante dessas correlações, as decisões baseadas em *Big Data* promovem maior acurácia do que originalmente, sendo esse o efeito econômico mais relevante. Por outro lado, o efeito social disso deve ser também verificado, principalmente quando há o risco de ocorrência de algum tipo de ato discriminatório com base na tomada de decisões usando *Big Data*.

3 – Discriminação algorítmica: conceito, tipologia e desafios jurídicos

Após identificar os principais conceitos ligados à tecnologia, cabe demonstrar que a utilização do *Big Data* ocorre nas relações de emprego, principalmente para gestão de pessoas, apresentando-se como uma das formas de busca mais comuns atualmente pela eficiência em diversas empresas, havendo de se verificar as repercussões disso. A utilização da IA para tomada de decisões, notadamente quando se tratando de gestão de pessoas, pode trazer aspectos discriminatórios, por vezes conscientes, mas geralmente inconscientes.

3.1 – Discriminação algorítmica e sua tipologia

Geralmente assume-se que a decisão que seja tomada por um mecanismo de IA será mais objetiva do que uma decisão tomada por um ser humano. Essa é a visão que pessoas comumente possuem em seu cotidiano, já que se imagina que uma máquina não terá emoções e, portanto, não decidirá tendendo para favorecer ou prejudicar alguém.

Reforçando o que foi dito, Laura Mendes e Marcela Mattiuzzo consideram que realmente os seres humanos são tendenciosos, ao passo em que o uso de algoritmos de IA, por exemplo, no recrutamento de novos empregados poderia reduzir os vieses discriminatórios. No entanto, um possível efeito adverso da utilização exclusiva dessa tecnologia poderia ocorrer, pois poderia racionalizar os critérios pretendidos pela empresa e considerar que os candidatos considerados "fora do padrão" não desempenhariam tão bem determinadas atividades, haja vista que suas capacidades seriam outras, diferentes dos demais colegas de setor (2019, p. 46).

Diversos casos de discriminação gerada por algoritmos ocorrem, sendo que nem todos acontecem de maneira uniforme, tanto que uns são causados por fatores diversos dos outros, ou então geram impactos de maneiras diferentes, razão pela qual a doutrina elaborou a tipologia da discriminação algorítmica, sendo sistematizada por Laura Mendes e Marcela Mattiuzzo em quatro principais tipos.

O primeiro tipo de discriminação algorítmica seria aquele gerado por um erro estatístico (Mendes; Mattiuzzo, 2019, p. 51-52). Assim, ocorre quando há algum erro de contabilização dos dados disponíveis, seja pela incorreção na base de dados, por exemplo dados incorretamente coletados, seja por problemas no código do algoritmo. Nessa situação, o erro decorre da programação que teria sido mal calibrada, fugindo de seu propósito pretendido.

As autoras ilustram esse tipo diante da grande quantidade de erros que ocorre nos casos de identificação de *score* de crédito nos Estados Unidos, afirmando, baseadas em parecer formulado pela FTC (*Federal Trade Commission*) ao Congresso daquele país, apresentando uma margem de erro entre 10% e 21%, dependendo da natureza do erro (Mendes; Mattiuzzo, 2019, p. 53).

O segundo tipo de discriminação por algoritmo é aquele gerado por uma generalização (Mendes; Mattiuzzo, 2019, p. 52), ou seja, todas as variáveis estão matematicamente corretas, mas por algum motivo o indivíduo é relacionado automaticamente a um grupo de outros indivíduos, a partir do que o algoritmo entende que ele possui as mesmas categorias do grupo no qual está inserido. Assim, a generalização ocorre, retirando a individualidade das características de uma pessoa, passando esta a ser analisada como um todo em seu grupo estatístico.

É certo que todo modelo é uma generalização e que nem toda generalização gera algum tipo de prejuízo. No entanto, há de se considerarem os casos em que a generalização possa afetar os direitos dos titulares de dados. Por exemplo, se o indivíduo possui boas condições financeiras, mas reside em um local em que a maioria das pessoas não as possuem, poderá ele (pela análise do seu CEP) ter algum tipo de prejuízo, diante da generalização feita com base nos dados possuídos dos demais indivíduos do grupo (Mendes; Mattiuzzo, 2019, p. 52-53).

Um terceiro tipo de discriminação diz respeito àquelas hipóteses em que, para que o algoritmo chegue a uma conclusão, deverá haver o tratamento de dados pessoais sensíveis do indivíduo (Mendes; Mattiuzzo, 2019, p. 52). A Lei Geral de Proteção de Dados (LGPD) classifica como dados pessoais sensíveis aqueles que expressam mais diretamente as informações mais íntimas do ser humano, por exemplo dados relativos à origem racial ou étnica, convicção religiosa, preferência política, filiação a sindicato, referentes à saúde ou biométricos[6].

Poder-se-ia questionar por que tais decisões podem levar à discriminação, mesmo que não haja risco aparente. Ocorre que esses dados sensíveis, além de serem mais protegidos pela lei, podem ser mascarados por outras informações (*proxies*) que vão indicar os dados sensíveis do indivíduo sem que ele saiba desse tratamento. Nesse sentido, grupos que são historicamente marginalizados (negros, mulheres, alguns grupos religiosos) podem permanecer marginalizados, pelo uso de informações que mascararam o tratamento de dados sensíveis (Mendes; Mattiuzzo, 2019, p. 54).

O quarto tipo de discriminação por algoritmo é aquele em que há limitação de direitos. Ou seja, quando os dados estão sendo tratados, pelo menos uma das finalidades do tratamento de dados vai gerar limitações a direitos que são do indivíduo (Mendes; Mattiuzzo, 2019, p. 52-53). Nesse caso, todos os cálculos estatísticos podem estar corretos, mas seu resultado será a limitação de um direito previamente estabelecido. Essa limitação não precisa ser ativa, pode ser também passiva.

Um exemplo apresentado pelas autoras ocorreu na Alemanha, quando autoridades verificaram incentivos econômicos para que os *bureaus* de crédito pudessem considerar o exercício do direito de consulta a esse *score* de crédito pelos cidadãos como um fator para sua redução, pois a mera consulta poderia ser considerada uma atitude típica de um indivíduo que não possui condições saudáveis de crédito. Na prática, os indivíduos tinham limitado seu direito de acessar sua nota de crédito, pois poderiam sofrer consequências discriminató-

6 Art. 5º Para os fins desta Lei, considera-se: [...] II – dado pessoal sensível: dado pessoal sobre origem racial ou étnica, convicção religiosa, opinião política, filiação a sindicato ou a organização de caráter religioso, filosófico ou político, dado referente à saúde ou à vida sexual, dado genético ou biométrico, quando vinculado a uma pessoa natural.

rias pelo algoritmo (redução da sua pontuação de crédito) (Mendes; Mattiuzzo, 2019, p. 54-55).

Diante dessa tipologia, fica claro que a discriminação não ocorre apenas porque o programador inseriu conscientemente algum viés discriminatório naquela aplicação de IA. Além disso, mesmo que uma empresa utilize um mecanismo de IA que produz discriminação, não necessariamente esse efeito será intencional, pois é muito comum que seja resultado de um dos tipos citados, alguma falha nos dados que foram processados pelo algoritmo, algum erro de programação, alguma generalização automaticamente realizada, que desconsidere outros aspectos, algum tratamento de dados sensíveis de maneira mascarada, ou limitando algum direito do indivíduo.

Por outro lado, não quer dizer que não haja responsabilidades ou mesmo salvaguardas que devam ser adotadas pelas empresas. Em realidade, quando há a utilização de alguma dessas tecnologias, deverá haver bastante cuidado e respeito às medidas de proteção da pessoa titular dos dados, garantindo que haja a devida diversificação nos critérios utilizados, a transparência desses critérios, além de uma periódica revisão dos resultados gerados pelos algoritmos.

Juridicamente, há diversas implicações relevantes que advêm da utilização de mecanismos de IA, principalmente quanto à transparência de seus critérios, haja vista a grande variedade de casos concretos em que se pôde verificar a ocorrência de discriminação algorítmica. Essas repercussões jurídicas serão, a seguir, abordadas.

3.2 – Implicações jurídicas da discriminação algorítmica

A literatura, nacional e internacional, apresenta diversos exemplos de discriminação gerada pelo uso de algoritmos de IA.

Cathy O'Neil é uma norte-americana, ativista de direitos humanos na *internet*, que acompanha, desde muito tempo, situações de mau uso de algoritmos de IA, considerando-os como armas de destruição matemática, comparando com armas de destruição em massa (utilizando o trocadilho na língua inglesa: "*weapons of math destruction*") (2016).

Um dos exemplos apresentados em seu livro é o do IMPACT, sistema baseado em algoritmo de avaliação de professores, que foi utilizado pelas escolas públicas de Washington, D.C. Esse sistema avaliava os professores de acordo com critérios que incluíam as notas dos alunos em testes ao final de cada ano escolar. Os resultados desse algoritmo geravam uma nota para os professores, os quais seriam dispensados caso tivessem uma nota baixa. Afinal, a intenção da escola era que o professor gerasse bons resultados nas notas de seus alunos.

Ocorre que, após diversos casos de descontentamento de professores, passou-se a investigar o algoritmo e percebeu-se que ele estava incentivando os professores a diminuírem o rigor em seus testes, já que isso provavelmente poderia impactar em sua continuidade no emprego. Mas, aqueles que não adotavam tais práticas, ou seja, aplicavam testes mais rigorosos a seus alunos, eram dispensados, pois a nota parecia refletir, de algum modo, as notas dos alunos.

O ponto que mais chama a atenção é que, quando um professor se sentia injustiçado, como a professora Sarah Bax, não haveria a quem recorrer. No caso dela, após recorrer a um superior, ficou sem resposta sobre uma pergunta simples: qual critério foi utilizado para aquela nota. Após tentativas de que técnicos pudessem esclarecer esses critérios, não foi possível, o que implicou o questionamento da autora: como é possível recorrer de uma decisão de cujo critério ninguém tem conhecimento? (O'Neil, 2016, p. 8).

Outro caso também citado reflete mais diretamente como os algoritmos podem ser discriminatórios, além de não serem transparentes (como no caso anteriormente citado).

A autora relata o caso de Kyle Behm, um jovem estadunidense que teve problemas em ser admitido em algumas seleções de emprego. Ele não conseguia entender o motivo dessas reprovações, se todas as suas notas nos exames SAT (algo como o vestibular brasileiro, mas nos Estados Unidos) eram elevadas. Ao reparar bem nos testes de personalidade de que ele participava nessas seleções, percebeu que se pareciam demais com um exame que havia feito, o qual o diagnosticou com transtorno bipolar.

Depois que descobriu que provavelmente estava sendo recusado em oportunidades de emprego por causa de sua condição, ajuizou ações contra sete empresas pelo uso desse teste (desenvolvido pela empresa Kronos) e contra a própria Kronos, por considerá-la responsável pela sua dificuldade de recolocação em emprego (O'Neil, 2016, p. 105-107).

Esses exemplos demonstram que alguns critérios utilizados em processos de automação, quando se utilizam mecanismos de IA, podem ser utilizados para buscar menores tendências pessoais ou reduzir vieses, porém esses próprios algoritmos de IA podem perpetuar discriminação baseada em vieses.

Esse tipo de situação ocorre, principalmente, pois os dispositivos de IA são tendencialmente obscuros, sem critérios de transparência que demonstrem os motivos de se chegar àquela decisão específica. Regra geral, algoritmos de aprendizado de máquina (*machine learning*) possuem essa dificuldade, pois o algoritmo passa a aprender com a própria prática decisória dele e com os dados a que é submetido. No entanto, essa base de dados pode estar enviesada.

A opacidade desses algoritmos é uma grande barreira à verificação da ocorrência de discriminação algorítmica. Segundo Danilo Doneda e colaboradores (2018, p. 6), a obscuridade do processo decisório do algoritmo é um grande problema para toda a sociedade, principalmente para as autoridades reguladoras, pois a verificação de uma possível discriminação somente será realizada se for possível saber qual o *input* ou qual o método estatístico utilizado.

Não à toa que os algoritmos são denominados, muitas vezes, de caixas pretas (*black boxes*) (Pasquale, 2015), pois, dependendo da metodologia utilizada para o aprendizado de máquina, será praticamente impossível verificar seus critérios ou sua árvore de decisão.

Dessa opacidade do algoritmo surgem diversos questionamentos sobre se eles deveriam incorporar em suas programações a transparência, principalmente considerando que é um dos princípios elencados pela LGPD[7].

Claramente, a exigência de transparência sempre que haja o uso de algoritmos de IA protegeria os direitos dos indivíduos. Mas alguns grandes problemas podem surgir desse tipo de pensamento amplo e genérico. Primeiramente, deve-se considerar que o proprietário do algoritmo o tem para diferenciar-se de seus concorrentes, pois poderá desenvolver inovação ao ponto de realizar algumas atividades que não sejam executáveis por outros. Assim, como garantir o segredo do negócio, se terá de ser plenamente transparente em relação a seu algoritmo de IA?

Inclusive, nesse sentido a LGPD flexibilizou o princípio da transparência (o qual já resguarda os segredos comercial e industrial), permitindo que o agente de tratamento que realize decisões automatizadas somente tenha de revelar informações sobre critérios caso seja solicitado pelo titular de dados. E, mesmo assim, poderá recusar-se a prestar maiores detalhes sobre a tecnologia, fundamentando nos segredos comercial e industrial[8]. No entanto, havendo a recusa, a Autoridade Nacional de Proteção de Dados poderá realizar auditoria, buscando exatamente verificar se haverá aspectos discriminatórios naquela decisão automatizada[9].

---

7 Art. 6º As atividades de tratamento de dados pessoais deverão observar a boa-fé e os seguintes princípios: [...] VI – transparência: garantia, aos titulares, de informações claras, precisas e facilmente acessíveis sobre a realização do tratamento e os respectivos agentes de tratamento, observados os segredos comercial e industrial.

8 Art. 20. O titular dos dados tem direito a solicitar a revisão de decisões tomadas unicamente com base em tratamento automatizado de dados pessoais que afetem seus interesses, incluídas as decisões destinadas a definir o seu perfil pessoal, profissional, de consumo e de crédito ou os aspectos de sua personalidade. § 1º O controlador deverá fornecer, sempre que solicitadas, informações claras e adequadas a respeito dos critérios e dos procedimentos utilizados para a decisão automatizada, observados os segredos comercial e industrial.

9 Art. 20. [...] § 2º Em caso de não oferecimento de informações de que trata o § 1º deste artigo baseado na observância de segredo comercial e industrial, a autoridade nacional poderá realizar auditoria para verificação de aspectos discriminatórios em tratamento automatizado de dados pessoais.

Mas o debate gira não apenas sobre a possibilidade jurídica de transparência, também quanto à possibilidade técnica, pois é discutível a transparência plena dos critérios de um algoritmo de IA.

Para Laura Mendes e Marcela Mattiuzzo (2019, p. 47), o algoritmo sendo obscuro seria bastante difícil verificar eventual discriminação que tenha ocorrido, sendo complicado também antecipar-se a possíveis discriminações.

Kroll e colaboradores (2017, p. 633) asseveram que a transparência não seria tão efetiva quanto aparenta ser, tendo em vista que poderá ser inatingível juridicamente, diante de algum argumento público ou razões privadas, bem como porque saber os métodos que foram adotados não seria suficiente se o algoritmo já não tivesse sido construído para permitir sua prestação de contas. Além disso, as aleatoriedades criadas pela IA seriam imprevisíveis em algumas metodologias, ainda mais quando se utiliza o *machine learning*.

Assim, mais do que exigir transparência, os autores sustentam que a prestação de contas poderia ser uma forma de garantir que a própria empresa deixe registrado seu tratamento de dados, para posterior auditoria ou verificação (Kroll *et al.*, 2017, p. 633).

4 – Inteligência Artificial na relação de emprego: eficiência econômica x fragilidade e nudez tecnológica do empregado

O uso de mecanismos tecnológicos no ambiente de trabalho é bastante comum, até porque faz parte da relação empregatícia estar sempre conectada com inovações. Essa ligação está, geralmente, associada ao interesse do empregador em reduzir custos de execução da atividade, atingindo cada vez mais eficiência.

Considerando, então, a busca pela eficiência no ambiente de trabalho, a seguir será mais bem delimitado como isso pode ocorrer, principalmente com o uso de mecanismos de Inteligência Artificial. Além disso, será analisado de que maneira a relação de emprego intensifica os efeitos danosos da discriminação gerada pelo uso dessas tecnologias algorítmicas.

4.1 – Eficiência e produtividade como objetivos da Inteligência Artificial no ambiente de trabalho

Um estudo realizado por Ekkehard Ernst (2018), para subsidiar a Comissão Global sobre o Futuro do Trabalho, da OIT, identificou três principais tipos de aplicação da IA que demandam maior atenção.

O primeiro desses usos é com a finalidade de combinação de resultados (*matching*). Por meio desses mecanismos de IA, a empresa consegue analisar

os dados que possui e traçar estratégias de atingimento de públicos específicos com as informações mais apropriadas para eles, sem que haja publicidade desnecessária para indivíduos que não possuam interesse em determinados temas.

O segundo uso frequente da IA no âmbito empresarial é com a finalidade de classificação. Assim, a empresa consegue, de maneira automatizada, identificar perfis de comportamento e encaixar determinados indivíduos nos perfis predefinidos, aplicando maior rigor na vigilância para conseguir acessar mais e mais desses dados.

O terceiro tipo de uso é aquele direcionado para a gestão de processos (*process-management*), quando os mecanismos de IA contribuem na eficiência de processos internos à empresa, tais como integração de cadeias de abastecimento, aprimoramento das informações sobre qualidade de produto, esquemas de certificação e condições de mercado.

Percebe-se que em todos os cenários de utilização empresarial da IA existe pertinência de essa utilização ocorrer, também, para objetivos internos.

Internalizando todas essas operações automatizadas, com uso de ferramentas de IA, o empregador consegue identificar aspectos relevantes para seu modelo de negócios a partir da gestão das pessoas, de maneira estratégica e utilizando modelos estatísticos.

Trata-se de uma abordagem bastante difundida, cuja principal denominação é *People Analytics* (ou *HR Analytics*).

Para Matthew Bodie, em seu artigo seminal "*The Law and Policy of People Analytics*", *People Analytics* é considerado um processo ou método de gestão de recursos humanos, baseado na utilização do *Big Data* para capturar informações relevantes a partir da *performance* laboral. A ideia chave do *People Analytics* é que julgamentos subjetivos não são tão rigorosos ou confiáveis quanto análises estatísticas sobre pessoas (Bodie, 2016, p. 3).

A abordagem *People Analytics* possui dois grandes objetivos, sendo o primeiro deles a competitividade organizacional, já que uma tomada de decisões baseada em dados estatísticos sobre resultados é muito mais efetiva na melhoria da gestão; o segundo objetivo é o desempenho e desenvolvimento individual, que exige um olhar para as pessoas buscando a otimização dos recursos humanos (Garrido; Silveira; Silveira, 2018, p. 35).

Geralmente essas análises estatísticas na gestão de pessoas são feitas com análises preditivas, ou seja, aquelas que buscam antecipar cenários a partir de algoritmos.

O *People Analytics* promove uma modificação radical na forma de a empresa enxergar seus problemas internos e resolvê-los, já que deixa de focar

nas causalidades (ou seja, quando as situações são previstas porque ocorreu uma outra situação, em relação de causa e efeito) e passa a focar nas correlações (quando duas ou mais situações ocorrem concomitantemente, mas não necessariamente são causa e efeito).

Asseveram Laura Mendes e Marcela Mattiuzzo (2019, p. 44) que:

> Durante muito tempo, a ciência e a descoberta científica funcionaram por meio da busca por causalidades. [...] Com *Big Data*, contudo, a causalidade perde espaço para as correlações. Uma correlação é a probabilidade de um evento ocorrer caso outro evento também se realize. É uma relação estatística entre tais acontecimentos.

Essas correlações são os grandes diferenciais das tecnologias baseadas em *Big Data*, pois conseguem, mais do que palpites, formular cenários probabilisticamente previsíveis. Esse tipo de predição não ocorre sem o uso de dados, afinal são necessários para toda e qualquer análise de cenários, principalmente por meio de algoritmos de IA.

Esses dados que serão tratados, visando a resultados econômicos, são os dados dos empregados, exatamente aqueles que possuem uma relação de subordinação com o empregador. Por isso, o questionamento sobre quais fragilidades podem estar presentes na relação de emprego, deixando o empregado mais suscetível à discriminação algorítmica, será respondido a seguir.

4.2 – Fragilidade e nudez tecnológica do empregado

Identificados os principais usos do *Big Data* e de mecanismos de IA nas relações de emprego, principalmente com a finalidade de atingir maior eficiência e produtividade no ambiente de trabalho, utilizando modelos de predição de comportamento, cabe, então, verificar quais possíveis peculiaridades poderiam levar maiores prejuízos para a relação de emprego, quando da utilização de mecanismos de IA pelo empregador, em comparação com outras relações jurídicas.

O ponto inicial de qualquer abordagem sobre uma relação de emprego implica considerar que se trata de uma relação desigual, em que uma parte possui muitos meios de tomada de decisão, enquanto a outra se encontra subordinada e limitada, além de prestar um serviço de maneira personalíssima.

Esse desequilíbrio na relação de emprego é reforçado mais ainda quando o empregador utiliza ferramentas tecnológicas, seja para rastrear condutas realizadas por seus empregados, seja para aferir produtividade, ou, ainda, para vigilância, até mesmo sobre aquilo que o empregado acessa nas dependências (físicas ou virtuais) da empresa.

Ressalte-se que esse desequilíbrio entre as partes é decorrente do poder empregatício, o qual não deve ser compreendido como um direito potestativo do empregador, ou mesmo como um direito subjetivo dele, muito menos como um mecanismo hierárquico na empresa, pois tais concepções não refletem as conformidades democráticas da empresa, vivenciadas a partir do século XX (Delgado, 1996, p. 185).

Para Mauricio Godinho Delgado, o poder intraempresarial (ou poder empregatício) deve ser considerado como uma relação jurídica contratual complexa, pois comporta a união de vontades (do empregador e do empregado), mas com variações de intensidade quanto ao peso dos indivíduos dessa relação, gerando uma assimetria variável, "mediante a qual se preveêm [sic], se alcançam ou sancionam condutas no plano do estabelecimento e da empresa" (Delgado, 1996, p. 191).

Essa concepção demonstra que o poder empregatício não se traduz como apenas uma relação de hierarquia, em que todos os comandos do empregador devem ser verticalmente aceitos sem qualquer participação do empregado, o qual também possui pretensões, liberdades e interesses naquela relação jurídica.

O poder empregatício legitima ao empregador as atividades ligadas à direção, à punição e ao controle. Essas atividades refletem as dimensões do poder empregatício, pois são componentes necessários para a boa gestão do empreendimento pelo empregador. Uma das dimensões do poder empregatício é o poder de controle, o qual é considerado por Alexandre Agra Belmonte como "prerrogativa de fiscalização da execução das tarefas conforme as diretrizes, métodos, metas e fins estabelecidos no contrato de trabalho" (Belmonte, 2004, p. 61).

É o poder de controle do empregador que o permite instalar câmeras de segurança no local de trabalho, restringir acesso a alguns *sites* pelos empregados, apresentar e monitorar alguns detalhamentos da rotina de trabalho que deve ser seguida, enfim, tudo aquilo que diz respeito à verificação do que está sendo realizado no ambiente de trabalho.

Apesar de facilitar o estudo sistemático do tema, a doutrina mais tradicional do Direito do Trabalho considera que o poder de controle não passa de uma manifestação do poder diretivo, também ligada ao poder disciplinar, do qual geralmente é o pressuposto (Delgado, 1996, p. 178). Ainda assim, não se deve confundir com o poder disciplinar, pois não necessariamente o controle será realizado com a finalidade única de disciplina do local de trabalho, mas, como visto na subseção 4.1 deste artigo, também buscando maior eficiência na execução das atividades, gerando maior produtividade.

O poder de controle, no entanto, tem a aptidão de trazer mais desequilíbrio ainda à relação de emprego, principalmente com a utilização das novas tecnologias do século XXI. Teresa Coelho Moreira considera que, com o uso das Novas Tecnologias da Informação e Comunicação (NTIC), esse poder de controle é atualizado, aumentado e, mesmo permanecendo como legítimo, poderá impactar direitos fundamentais do empregado, sobretudo aqueles mais ligados a ele enquanto pessoa (Moreira, 2010, p. 32-33), ou seja, os direitos de personalidade.

Portanto, a própria existência de um legítimo poder de controle na relação de emprego, que já permitiria a vigilância ser executada no local de trabalho, diferencia a relação de emprego das demais relações jurídicas.

Além da própria existência de um legítimo poder de controle, esse mesmo controle é exercido de uma maneira diferenciada na relação de emprego. A vigilância no local de trabalho, principalmente diante de uma realidade digital, ocorre de maneira imperceptível, oculta (Moreira, 2010, p. 62).

Essa característica da vigilância laboral, obtida por meio das novas tecnologias, reflete o que se denomina hoje de um novo poder de controle, pois remodela a relação de vigilância tradicional. Enquanto tradicionalmente o empregador utilizava sua estrutura hierárquica para impor algum tipo de controle centralizado (revista íntima ou monitoramento por câmeras), com as novas tecnologias, o controle se dá de maneira muito mais sutil e muito mais descentralizada.

Volte-se ao caso citado na subseção 3.2, em que o algoritmo desenvolvido pela empresa Kronos estaria aplicando testes de personalidade, aparentemente inofensivos, mas que na realidade estavam identificando pessoas que possuíssem algum tipo de doença ou mesmo condição mental fora do padrão.

Acrescente-se, ainda, outro exemplo emblemático citado por Cathy O'Neil (2016, p. 132-133), envolvendo a empresa americana Cataphora, que, em 2008, comercializou um *software* que classificava empregados de acordo com uma série de métricas, dentre as quais sua geração de ideias. Na prática, o *software* acessava os *e-mails* e as mensagens corporativas dos empregados, classificando-os de acordo com sua capacidade de criar ideias ou de conectá-las.

O problema, segundo a autora, é que as empresas podem preferir demitir empregados que tenham menores classificações nesses critérios (presumindo que tais empregados seriam menos criativos), porém eles poderiam estar gerando excelentes ideias sem compartilhá-las com os colegas, ou mesmo poderiam ter contribuição em outros aspectos, por exemplo melhorando o clima corporativo com piadas ou comentários úteis aos demais, características igualmente valorizadas no ambiente de trabalho (O'Neil, 2016, p. 132-133).

Esse tipo de inferência sobre os empregados ocorre visando a melhores resultados, selecionando aqueles empregados mais ajustados aos critérios pretendidos pela empresa, porém podem gerar impactos discriminatórios, pois esses empregados não podem sequer saber que estão sendo objeto dessas inferências, muito menos de maneira automatizada, ou seja, com uso de Inteligência Artificial.

Então, diante do legítimo poder de controle, atualizado pelas novas tecnologias, tem-se um monitoramento eletrônico ocorrendo nas relações de emprego, que acontece de maneira imperceptível, sutil e descentralizada, fragilizando o empregado nesse tipo de situação.

Para além da fragilidade própria a esse tipo de relação jurídica, outro critério é capaz de diferenciar as relações de emprego das demais. Ao ser contratado, um empregado já precisa fornecer diversos dados pessoais a seu empregador, seja por meio de envio de documentos próprios, seja por meio do preenchimento de algum tipo de formulário (físico ou digital).

Esses dados pessoais são utilizados para diversas finalidades, sendo a principal delas permitir que haja o poder diretivo do empregador, que ele possa identificar seus empregados, direcioná-los para o exercício de suas atividades, prestar contas de suas obrigações legais ou sindicais, portanto, para fazer com que o contrato de trabalho se efetive.

No entanto, esses não são os únicos dados acessíveis pelo empregador, principalmente quando são utilizadas novas tecnologias, a exemplo de algoritmos de IA. O empregador tem acesso às trocas de mensagens dos empregados, quando utilizando alguma ferramenta de comunicação interna, a registros de horários dos empregados, algumas vezes a dados de geolocalização, dentre diversos outros.

Ao buscar uma gestão de pessoas mais efetiva, o empregador terá de inferir algumas conclusões a partir dos dados dos empregados, sendo que os dados básicos não serão suficientes. O empregador poderá ter acesso a dados pessoais sensíveis dos empregados, àqueles dados ligados à saúde, ou mesmo àquelas informações dos familiares do empregado.

Todo o acesso do empregador a essas informações, não apenas àquelas estritamente necessárias para a finalidade principal do contrato de trabalho, mas que são importantes para a melhoria da gestão empresarial, demonstra que o empregado está cada dia mais desnudado pelo empregador. A cada dia mais o empregador tem acesso a quem seu empregado é, faz ou deixa de fazer.

Teresa Coelho Moreira afirma que se trata de uma nudez tecnológica do empregado, pois essas novas tecnologias aumentam desmesuradamente a capa-

cidade de acumulação de informações, permitindo criar um perfil do empregado com base em informações aparentemente inócuas e inofensivas (2010, p. 27).

Mais especificamente quanto à Inteligência Artificial, algumas peculiaridades são apontadas pela doutrina como especificamente sensíveis para as relações de emprego. Antonio Aloisi e Elena Gramano (2019, p. 105-106) sistematizam essas peculiaridades, listando-as em quatro principais.

Primeiro que esses mecanismos de IA geralmente coletam informações de diferentes fontes (mesmo fora da relação de emprego), além de coletar mais dados do que estritamente necessários. Além disso, esses algoritmos criam inferências não intuitivas e não verificáveis, gerando um *HR Analytics* baseado na predição. Ainda, os algoritmos podem causar impacto não só na liberdade e na privacidade dos empregados, mas também na sua autonomia e nas decisões morais, conectando cada vez mais a vida pessoal com a vida profissional. Por fim, esses mecanismos, associados a uma cultura de automonitoramento e gamificação, podem gerar o fornecimento de dados pessoais pelo empregado de maneira consensual, mas em troca de pequenos benefícios supérfluos.

Os impactos desse tipo de tratamento de dados, de maneira automatizada, pelo empregador são pouco discutidos e, menos ainda, regulados.

No Brasil, a LGPD traz um regramento sobre decisões automatizadas, porém permite que a empresa deixe de proativamente apresentar os critérios para a tomada de decisão automatizada, devendo aguardar a solicitação do titular de dados.

Especificamente na relação de emprego, não há legislação que trate do tema, abrindo a possibilidade de que o empregador utilize mecanismos de IA, sem que o empregado saiba que está sendo monitorado por eles, quais as finalidades daquele monitoramento ou, ainda, quais as consequências resultantes.

5 – Considerações finais

A busca pela eficiência na gestão empresarial é um valor inegável e irrecusável, haja vista que a própria concepção da empresa e da divisão do trabalho gira em torno desse objetivo. Diante disso, diversas tecnologias são aplicadas nas relações de emprego, principalmente visando a alcançar maior efetividade nas atividades desempenhadas por cada um dos empregados.

Mecanismos baseados em Inteligência Artificial são utilizados também nessas situações, principalmente em conjunto com o *Big Data*, que fornece uma imensa quantidade de dados, de uma maneira muito rápida e organizados de uma determinada forma que são dotados de um imenso valor.

Ainda que sejam utilizadas com vistas a finalidades empresariais legítimas, podem gerar decisões com caráter discriminatório. Essa discriminação gerada por um algoritmo possui diversas repercussões jurídicas, principalmente porque fazem com que aqueles que utilizam esses algoritmos devam adotar medidas de governança, garantindo transparência e prestando contas dos critérios nos quais aquelas decisões estariam baseadas.

Especificamente na relação de emprego, os algoritmos de IA são aplicados, por exemplo, para a gestão de pessoas de maneira automatizada, estatisticamente mais assertiva e buscando identificar as características de um indivíduo que podem ser mais adequadas a cada atividade. Ocorre que, nesse tipo de relação jurídica, os impactos discriminatórios da IA podem ser mais gravosos, diante da desigualdade entre as partes na relação de emprego, da fragilidade do empregado, mediante o legítimo poder de controle do empregador; e da nudez tecnológica, já que as tecnologias utilizadas pelo empregado coletam grandes quantidades de dados pessoais, muitas vezes desnecessários.

Essas peculiaridades das relações de emprego são pouco abordadas pela doutrina, assim como pela legislação, tanto que não há qualquer referência a isso na legislação trabalhista, apenas parâmetros gerais na LGPD, mas que esbarram nesses aspectos peculiares, deixando de ser tão efetivos. Em razão disso, haveria de ser mais bem delineada essa aplicação de IA na relação de emprego, por meio de legislação própria, que traga parâmetros específicos e adequados para a relação de emprego.

Referências

ALOISI, Antonio; GRAMANO, Elena. Artificial intelligence is watching you at work. Digital surveillance, employee monitoring and regulatory issues in the EU context. *Special Issue of Comparative Labor Law & Policy Journal, "Automation, Artificial Intelligence and Labour Protection"*, 2019, p. 105-106. Disponível em: http://salus.adapt.it/wp-content/uploads/2020/07/Gramano-Alois_AI-is-Wathching-you_2019.pdf. Acesso em: 3 set. 2021.

BELMONTE, Alexandre Agra. *O monitoramento da correspondência eletrônica nas relações de trabalho*. São Paulo: LTr, 2004.

BODIE, Matthew T. *et al*. The law and policy of people analytics. *University of Colorado Law Review*, v. 88, 2016, p. 3. Disponível em: https://scholarship.law.slu.edu/cgi/viewcontent.cgi?article=1001&context=faculty. Acesso em: 3 set. 2021.

DELGADO, Mauricio Godinho. *O poder empregatício*. São Paulo: LTr, 1996.

DOMINGOS, Pedro. *O algoritmo mestre*: como a busca pelo algoritmo de *machine learning* definitivo recriará nosso mundo. São Paulo: Novatec, 2017.

DONEDA, Danilo Cesar Maganhoto *et al*. Considerações iniciais sobre inteligência artificial, ética e autonomia pessoal. *Pensar – Revista de Ciências Jurídicas*, v. 23, n. 4, p. 1-17, 2018.

ERNST, Ekkehard *et al*. The economics of artificial intelligence: implications for the future of work, *ILO Future of Work Research Paper Series*; ILO, 2018.

GARRIDO, Giovanna; SILVEIRA, Rafael Damasco; SILVEIRA, Marco Antonio. *People analytics*: uma abordagem estratégica para a gestão do capital humano. *Revista Eletrônica de Estratégia & Negócios*, v. 11, n. 1, p. 28-52, 2018.

KAPLAN, Andreas; HAENLEIN, Michael. Siri, Siri, in my hand: who's the fairest in the land? On interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, [s.l.], v. 62, n. 1, p. 15-25, jan. 2018. Elsevier BV. Disponível em: http://dx.doi.org/10.1016/j.bushor.2018.08.004. Acesso em: 2 set. 2021.

KROLL, Joshua Alexander *et al*. Accountable algorithms. *University of Pennsylvania Law Review*, v. 165, p. 633-705, 2017.

MAYER-SCHÖNBERGER, V.; CUKIER, K. *Big Data*: a revolution that will transform how we live, work, and think. New York: First Mariner Books, 2014.

MENDES, Laura Schertel; MATTIUZZO, Marcela. Discriminação algorítmica: conceito, fundamento legal e tipologia. *Direito Público*, v. 16, n. 90, 2019.

MOREIRA, Teresa Coelho. *A privacidade dos trabalhadores e as novas tecnologias de informação e comunicação*: contributo para um estudo dos limites do poder de controlo electrónico do empregador. Coimbra: Almedina, 2010.

O'NEIL, Cathy. *Weapons of math destruction*: how big data increases inequality and threatens democracy. New York: Crown, 2016.

PASQUALE, Frank. *The black box society*. Cambridge: Harvard University Press, 2015.

SMITH, Adam. *A riqueza das nações*. 3. ed. São Paulo: WMF Martins Fontes, 2016. v. I.

TURING, Alan M.; HAUGELAND, J. *Computing machinery and intelligence*. Cambridge, MA: MIT Press, 1950.

Como citar este artigo:
MATIAS, João Luís Nogueira; MORAIS JÚNIOR, Ricardo Antônio Maia de. Discriminação algorítmica na relação de emprego: eficiência econômica, inteligência artificial e fragilidade do empregado. *Revista do Tribunal Superior do Trabalho*, Porto Alegre, v. 90, n. 2, p. 128-147, abr./jun. 2024.