

ano 22 – n. 90 | outubro/dezembro – 2022

Belo Horizonte | p. 1-236 | ISSN 1516-3210 | DOI: 10.21056/aec.v22i90

A&C – R. de Dir. Administrativo & Constitucional

www.revistaaec.com

A&C

**Revista de Direito
ADMINISTRATIVO
& CONSTITUCIONAL**

**A&C – ADMINISTRATIVE &
CONSTITUTIONAL LAW REVIEW**

FORUM

A246 A&C : Revista de Direito Administrativo & Constitucional. – ano 3, n. 11, (jan./mar. 2003). – Belo Horizonte: Fórum, 2003-

Trimestral

ISSN impresso 1516-3210

ISSN digital 1984-4182

Ano 1, n. 1, 1999 até ano 2, n. 10, 2002 publicada pela Editora Juruá em Curitiba

1. Direito administrativo. 2. Direito constitucional.
I. Fórum.

CDD: 342

CDU: 342.9

Coordenação editorial: Leonardo Eustáquio Siqueira Araújo
Aline Sobreira de Oliveira

Capa: Igor Jamur

Projeto gráfico: Walter Santos

Periódico classificado no Estrato A3 do Sistema Qualis da CAPES - Área: Direito.

Qualis – CAPES (Área de Direito)

Na avaliação realizada em 2022, a revista foi classificada no estrato A3 no Qualis da CAPES (Área de Direito).

Entidade promotora

A A&C – *Revista de Direito Administrativo & Constitucional*, é um periódico científico promovido pelo Instituto de Direito Romeu Felipe Bacellar com o apoio do Instituto Paranaense de Direito Administrativo (IPDA).

Foco, Escopo e Público-Alvo

Foi fundada em 1999, teve seus primeiros 10 números editorados pela Juruá Editora, e desde o número 11 até os dias atuais é editorada e publicada pela Editora Fórum, tanto em versão impressa quanto em versão digital, sediada na BID – Biblioteca Digital Fórum. Tem como principal objetivo a divulgação de pesquisas sobre temas atuais na área do Direito Administrativo e Constitucional, voltada ao público de pesquisadores da área jurídica, de graduação e pós-graduação, e aos profissionais do Direito.

Linha Editorial

A linha editorial da A&C – *Revista de Direito Administrativo & Constitucional*, estabelecida pelo seu Conselho Editorial composto por renomados juristas brasileiros e estrangeiros, está voltada às pesquisas desenvolvidas na área de Direito Constitucional e de Direito Administrativo, com foco na questão da efetividade dos seus institutos não só no Brasil como no Direito comparado, enfatizando o campo de interseção entre Administração Pública e Constituição e a análise crítica das inovações em matéria de Direito Público, notadamente na América Latina e países europeus de cultura latina.

Cobertura Temática

A cobertura temática da revista, de acordo com a classificação do CNPq, abrange as seguintes áreas:

- Grande área: Ciências Sociais Aplicadas (6.00.00.00-7) / Área: Direito (6.01.00.00-1) / Subárea: Teoria do Direito (6.01.01.00-8) / Especialidade: Teoria do Estado (6.01.01.03-2).
- Grande área: Ciências Sociais Aplicadas (6.00.00.00-7) / Área: Direito (6.01.00.00-1) / Subárea: Direito Público (6.01.02.00-4) / Especialidade: Direito Constitucional (6.01.02.05-5).
- Grande área: Ciências Sociais Aplicadas (6.00.00.00-7) / Área: Direito (6.01.00.00-1) / Subárea: Direito Público (6.01.02.00-4) / Especialidade: Direito Administrativo (6.01.02.06-3).

Indexação em Bases de Dados e Fontes de Informação

Esta publicação está indexada em:

- Web of Science (ESCI)
- Ulrich's Periodicals Directory
- Latindex
- Directory of Research Journals Indexing
- Universal Impact Factor
- CrossRef
- Google Scholar
- RVBI (Rede Virtual de Bibliotecas – Congresso Nacional)
- Library of Congress (Biblioteca do Congresso dos EUA)
- MIAR - Information Matrix for the Analysis of Journals
- WorldCat
- BASE - Bielefeld Academic Search Engine
- REDIB - Red Iberoamericana de Innovación y Conocimiento Científico
- ERIHPLUS - European Reference Index for the Humanities and the Social Sciences
- EZB - Electronic Journals Library
- CiteFactor
- Diadorim

Processo de Avaliação pelos Pares (Double Blind Peer Review)

A publicação dos artigos submete-se ao procedimento *double blind peer review*. Após uma primeira avaliação realizada pelos Editores Acadêmicos responsáveis quanto à adequação do artigo à linha editorial e às normas de publicação da revista, os trabalhos são remetidos sem identificação de autoria a dois pareceristas *ad hoc* portadores de título de Doutor, todos eles exógenos à Instituição e ao Estado do Paraná. Os pareceristas são sempre Professores Doutores afiliados a renomadas instituições de ensino superior nacionais e estrangeiras.

O dilema ético da decisão algorítmica na administração pública

The ethical dilemma of algorithmic decision in Public Administration

André Felipe Silva Puschel*

Universidade de São Paulo (Brasil)
andre_puschel@hotmail.com
<https://orcid.org/0000-0002-5705-7986>

Roberto Tessis Rodrigues**

Pontifícia Universidade Católica do Paraná (Brasil)
tessisroberto@gmail.com
<https://orcid.org/0000-0002-9499-4004>

Vivian Cristina Lima López Valle***

Pontifícia Universidade Católica do Paraná (Brasil)
vivian.lima@pucpr.br
<https://orcid.org/0000-0002-5793-2912>

Recebido/Received: 12.01.2022/January 12th, 2022.
Aprovado/Approved: 17.09.2022/September 17th, 2022.

Resumo: A utilização da inteligência artificial como suporte à tomada de decisão é realidade no cenário atual. Discussões sobre processamento de dados – sejam estas por meio de aprendizagem supervisionada, sejam por aprendizagem não supervisionada (*deep learning*) – ganham relevância no âmbito do direito. Ao mesmo tempo em que ganham relevância, porém, surgem os desafios sobre as

Como citar este artigo/*How to cite this article*: PUSCHEL, André Felipe Silva; RODRIGUES, Roberto Tessis; VALLE, Vivian Cristina Lima López. O dilema ético da decisão algorítmica na administração pública. *A&C – Revista de Direito Administrativo & Constitucional*, Belo Horizonte, ano 22, n. 90, p. 207-226, out./dez. 2022. DOI: 10.21056/aec.v22i90.1737.

* Mestrando em Direito pela Universidade de São Paulo (São Paulo-SP, Brasil). Bacharel em Direito pela Pontifícia Universidade Católica do Paraná. *E-mail*: andre_puschel@hotmail.com.

** Bacharel em Direito pela Pontifícia Universidade Católica do Paraná (Curitiba-PR, Brasil). *E-mail*: tessisroberto@gmail.com.

*** Professora Titular do Curso de Direito da Pontifícia Universidade Católica do Paraná (Curitiba-PR, Brasil). Doutora e Mestre em Direito do Estado pela Universidade Federal do Paraná. *E-mail*: vivian.lima@pucpr.br.

resultantes que se apresentam em razão do processamento de dados (*output*), na medida em que estes podem estar enviesados em função do conjunto de dados (*dataset*) provenientes do mundo físico, das concepções e das relações humanas. Por conta desses desafios, o presente artigo tem por finalidade pontuar tais aspectos no contexto contemporâneo e propor medidas que possam ser adotadas como possíveis soluções. A análise acerca da existência de um antídoto como contraposição ao enviesamento faz-se necessária no intuito de qualificar o uso contínuo e irrefreável dessas tecnologias. Tratar as dificuldades do tema no âmbito constitucional e à luz da legislação brasileira, por sua vez, justifica a importância do assunto e permite uma proposição para tais desafios, que se apresentam em razão do uso da inteligência artificial. Diante disso, o presente artigo utiliza-se da pesquisa bibliográfica sobre o assunto e de interpretação sistêmica entre a Constituição Federal e as legislações que abordam aspectos legais sobre a administração pública.

Palavras-chave: Inteligência artificial. Algoritmos. Decisão algorítmica. *Deep learning*. Processamento de dados.

Abstract: The use of artificial intelligence to support decision-making is a reality in the current scenario. Discussions on data processing – whether through supervised learning or through unsupervised learning (deep learning) – gain relevance in the field of law. At the same time that they gain relevance, however, challenges arise regarding the results that emerge due to data processing (output), insofar as these may be biased due to the data set coming from the physical world, of conceptions and human relations. Due to these challenges, this article aims to point out such aspects in the contemporary context and propose measures that can be adopted as possible solutions. The analysis about the existence of an antidote as a counterpoint to the bias is necessary, in order to qualify the continuous and unstoppable use of these technologies. Dealing with the difficulties of the subject in the constitutional scope and in the light of Brazilian legislation, in turn, justifies the importance of the subject and allowed a proposal for such challenges that arise due to the use of artificial intelligence. Therefore, this article makes use of bibliographical research on the subject and of systemic interpretation between the Federal Constitution and legislation that address legal aspects of Public Administration.

Keywords: Artificial intelligence. Algorithms. Algorithmic decision. Deep learning. Data processing.

Sumário: Introdução – **1** Técnicas de automação da aprendizagem – **2** O funcionamento da tomada de decisão e o dilema dos algoritmos decisórios – **3** A compreensão dos vieses decisórios algorítmicos – **4** A compreensão atual do sistema brasileiro e seus aspectos deserviesantes para superar o dilema ético da decisão algorítmica – **5** Conclusão – Referências

Introdução

A inteligência artificial tem modificado a sociedade, sobretudo a partir da constante evolução na capacidade de processamento e armazenamento de dados (*big data*).¹ As interações entre o homem e a inteligência artificial, por conta disso, têm sido cada vez mais intensas e mostram muitos benefícios à sociedade em geral;² dentre esses, verificam-se: a) a facilidade com que pessoas realizam, de

¹ Como citar este artigo/ *How to cite this article*: [inserir referência completa nas normas da ABNT – LEMBRAR AO FINAL DE INSERIR PÁGINA INICIAL E PÁGINA FINAL]. DOI:

BELLOCCHIO, Lucía; SANTIAGO, Alfonso. Estado digital de Derecho. *A&C – Revista de Direito Administrativo & Constitucional*, Belo Horizonte, ano 20, n. 80, p. 87-102, abr./jun. 2020.

² Sobre o tema, ver: NOHARA, Irene Patrícia; COLOMBO, Bruna Armonas. Tecnologias cívicas na interface entre direito e inteligência artificial: Operação Serenata de Amor para gostosuras ou travessuras? *A&C – Revista de Direito Administrativo & Constitucional*, Belo Horizonte, ano 19, n. 76, p. 83-103, abr./jun. 2019; VALLE, Vivian Cristina Lima López; GALLO, William Ivan. Inteligência artificial e capacidades regulatórias do Estado no ambiente da administração pública digital. *A&C – Revista de Direito Administrativo & Constitucional*,

qualquer lugar, reuniões *online* com outras em diversas localidades do mundo; b) a automação de processamento de informações utilizada por empresas como Amazon, YouTube, Netflix, entre outras, que identificam os conteúdos (ou produtos) mais utilizados (ou pesquisados) pelo usuário e sugerem conteúdos similares no intuito de converter novo consumo; c) aplicação em *smart regulation* por meio da Análise de Impacto Regulatório (RIA),³ que, com o uso da IA, analisa dados e informações encontrados à disposição do gestor público com objetivo de prever e evitar falhas de mercado e de governança.

Nesse sentido, a interação entre as pessoas e a inteligência artificial faz parte do cotidiano e tem sido objeto de constantes evoluções,⁴ principalmente quando nos deparamos com a realidade das decisões apoiadas pela inteligência artificial no âmbito da sociedade civil, nas relações comerciais, no Poder Judiciário, na esfera administrativa,⁵ no Poder Legislativo,⁶ entre outros. Contudo, considerando a forma com que as informações são processadas pelos sistemas, as suas resultantes podem apresentar muitas dúvidas em relação à confiança e à segurança jurídica dos dados apresentados, inclusive quando estamos diante da necessidade de atuação transparente e imparcial.

Desse modo, o cerne do presente artigo consiste na verificação desse desafio imposto à sociedade sobre o dilema ético da decisão algorítmica, de modo a verificar as dificuldades existentes no processamento das informações e propor possível solução, mediante interpretação sistêmica, para que se obtenha uma decisão algorítmica ética, imparcial e transparente.

Para tanto, serão apresentadas as técnicas de automação da aprendizagem, haja vista a necessidade de situar o leitor quanto às maneiras com que a inteligência

Belo Horizonte, ano 20, n. 82, p. 67-86, out./dez. 2020; BRAVO, Álvaro Avelino Sánchez. Marco Europeo para una inteligencia artificial basada en las personas. *International Journal of Digital Law*, Curitiba, v. 1, n. 1, p. 65-77, jan./abr. 2020.

³ GALLO APONTE, W. I.; LÓPEZ VALLE, V. C.; FÁCIO, R. N. La utilización de inteligencia artificial en la actividad regulatoria: una propuesta en favor del desarrollo nacional sostenible. *Veredas do Direito*, Belo Horizonte, v. 17, n. 39, p. 123-146, sep./dic. 2020.

⁴ FORNASIER, Mateus de Oliveira. The applicability of the Internet of Things (IoT) between fundamental rights to health and to privacy. *Revista de Investigações Constitucionais*, Curitiba, vol. 6, n. 2, p. 297-321, maio/ago. 2019.

⁵ STRINGHINI, Antonella. Administración Pública Inteligente: novedades al ecosistema normativo digital de la República Argentina. *Revista Eurolatinoamericana de Derecho Administrativo*, Santa Fe, vol. 5, n. 2, p. 199-215, jul./dic. 2018; STRINGHINI, Antonella. Asistencia virtual automatizada e inclusiva para optimizar la relación de la ciudadanía con la Administración Pública. *International Journal of Digital Law*, Curitiba, v. 1, n. 1, p. 117-128, jan./abr. 2020; VIANA, Ana Cristina Aguilar. Transformação digital na administração pública: do governo eletrônico ao governo digital. *Revista Eurolatinoamericana de Derecho Administrativo*, Santa Fe, vol. 8, n. 1, p. 115-136, ene./jun. 2021.

⁶ VALLE, Vivian Cristina Lima López; MARTINELLI, Gustavo; BARBOSA FILHO, Claudio Roberto. Participação no processo legislativo como direito fundamental: um olhar sobre as ferramentas de e-democracy no Congresso Nacional Brasileiro. *Revista Eurolatinoamericana de Derecho Administrativo*, Santa Fe, vol. 7, n. 1, p. 143-166, ene./jun. 2020.

artificial opera em relação ao processamento de informações e aos resultantes obtidos.

Por sua vez, será necessária a abordagem sobre a forma do funcionamento da tomada de decisão e o dilema dos algoritmos decisórios, ocasião em que serão apresentados os diversos cenários de riscos que podem surgir diante dessa automação da aprendizagem.

A pretensão, após a compreensão pelo leitor das técnicas de automação da aprendizagem e do funcionamento da tomada de decisão, consiste na análise dos motivos que potencializam os vieses humanos, de forma a delimitar a existência de um desafio quando da tomada de decisão. Isso porque a existência de vieses humanos, mediante a identificação de pré-conceitos existentes, assim como da possibilidade da própria inteligência artificial potencializar informações de acordo com o processamento (*machine learning*),⁷ enseja um desafio a ser superado por meio de um antídoto para tal.

Sendo assim, com a conseqüente delimitação do desafio e da proposição de um antídoto que pretende solucioná-lo, o artigo tem a finalidade de verificar o cenário atual brasileiro, bem como demonstrar que a importância do desafio a ser solucionado decorre do fato do tema consistir em direito fundamental e exigir seriedade, o que justifica a adoção de uma interpretação sistêmica para sua solução.

1 Técnicas de automação da aprendizagem

A integração da inteligência artificial (IA) como suporte à tomada de decisão está presente desde as iniciativas mercantis da sociedade civil, no Poder Judiciário, na administração pública, até as estruturas de controle, a exemplo do Tribunal de Contas da União.⁸ A despeito da utilização sistemática em diversos campos e dimensões da sociedade contemporânea de sistemas de tomada de decisão apoiada baseados em IA, o *design* dessas novas tecnologias, assim como a estrutura lógica ou a técnica utilizada para a tomada de decisão, permanece opaco. Esse gris deve-se ao fato de que a automação por meio de algoritmos envolve múltiplos processos, dentre eles a coleta, o processamento de dados e a execução das decisões com pouca ou nenhuma intervenção humana.

⁷ FIGUEIREDO, Carla Regina Bortolaz de; CABRAL, Flávio Garcia. Inteligência artificial: machine learning na Administração Pública. *International Journal of Digital Law*, Curitiba, v. 1, n. 1, p. 79-96, jan./abr. 2020.

⁸ A utilização de tecnologia como ferramenta de tratamento de informações pelo TCU vem de longa data, com iniciativas de aplicação de mecanismos cognitivos (SILVA, L. A. D. Uso de técnicas de inteligência artificial para subsidiar ações de controle. *Revista do TCU*, Brasília, n. 137, p. 124-129, set./dez. 2016). A iniciativa não se limita ao Tribunal de Contas da União e se vê disseminada por várias outras estruturas de controle (VALENTE, J. Órgãos públicos usam inteligência artificial para combater corrupção. *Agência Brasil*, 3 ago. 2018. Disponível em: <https://agenciabrasil.ebc.com.br/geral/noticia/2018-08/orgaos-publicos-usam-inteligencia-artificial-para-combater-corrupcao>. Acesso em: 1^o mar. 2020).

Diante do exposto, faz-se necessário compreender determinados termos e esclarecer a relação entre esses conceitos, sendo, respectivamente: inteligência artificial, algoritmo, aprendizado de máquina e decisão da máquina. Juarez Freitas⁹ define IA como sendo “um sistema algorítmico adaptável e relativamente autônomo, emulatório da decisão humana”; algoritmos, no livro *Justiça Digital*,¹⁰ é apresentado como “uma sequência de instruções que diz ao computador o que fazer”; aprendizado de máquina é a instrumentalização da IA no intuito de reproduzir a operação cognitiva e/ou decisória que se desenvolveria no cérebro humano;¹¹ e decisão da máquina é uma resultante do processo de aprendizagem.

Em síntese simplificada, os dados entram em um sistema (processo chamado de *input*), são processados pelo algoritmo, e o resultado desse processamento (*output*) sai do sistema. Nessa breve exemplificação, percebe-se que o resultante (*output*) é a decisão da máquina, portanto, a qualificação desse resultado contemplará eventual análise sobre o *input* (entrada de dados) e sobre a natureza do algoritmo que norteou o processamento dessa informação.

Os algoritmos podem ser programados ou não programados:¹² os primeiros terão definidos pelo programador todos os processos até que a informação saia do sistema; os segundos – objeto de nossos estudos – são os algoritmos de aprendizagem, que, por sua vez, podem ser de aprendizagem supervisionada, aprendizagem não supervisionada ou aprendizagem por reforço.¹³ Os algoritmos não programados – *learners* – são os que empregam o *machine learning* (aprendizagem de máquina).¹⁴

Na aprendizagem supervisionada de máquina, o *input* é selecionado pelo agente externo, e o *output* é uma resultante binária ou uma saída de multiclassificação.¹⁵ A título de clássico exemplo de saída binária, destaca-se a aprendizagem supervisionada dos filtros de *spam*, em que o usuário marca as mensagens que classifica como

⁹ FREITAS, Juarez; FREITAS, Thomas Bellini. *Direito e inteligência Artificial*. Belo Horizonte: Fórum, 2020. p. 13.

¹⁰ FERRARI, I. *Justiça digital* [livro eletrônico]. 2. ed. São Paulo: Thomson Reuters Brasil, 2021. RB-6.2.

¹¹ VALLE, Vanice Lírio do. Inteligência artificial incorporada à Administração Pública: mitos e desafios teóricos. *A&C – Revista de Direito Administrativo & Constitucional*, Belo Horizonte, ano 20, n. 81, p. 179-200, jul./set. 2020. p. 7.

¹² ARAÚJO, Valter Shuenquener de; ZULLO, Bruno Almeida; TORRES, Maurílio. Big Data, algoritmos e inteligência artificial na Administração Pública: reflexões para a sua utilização em um ambiente democrático. *A&C – Revista de Direito Administrativo & Constitucional*, Belo Horizonte, ano 20, n. 80, p. 241-261, abr./jun. 2020; CORVALÁN, Juan Gustavo. Inteligencia Artificial GPT-3, Pretoria y Oráculos Algorítmicos en el Derecho. *International Journal of Digital Law*, Curitiba, v. 1, n. 1, p. 11-52, jan./abr. 2020.

¹³ AMARILES, D. R. Algorithmic Decision Systems Automation and Machine Learning in the Public Administration. In: *The Cambridge Handbook of The Law of Algorithms*. Cambridge: Cambridge University Press, 2021. p. 305.

¹⁴ FERRARI, I. *Justiça digital* [livro eletrônico]. 2. ed. São Paulo: Thomson Reuters Brasil, 2021. RB-6.2.

¹⁵ AMARILES, D. R. Algorithmic Decision Systems Automation and Machine Learning in the Public Administration. In: *The Cambridge Handbook of The Law of Algorithms*. Cambridge: Cambridge University Press, 2021. p. 306.

spam e, com o tempo, a máquina estabelece os padrões das mensagens que detêm essa marcação e passa a decidir quais mensagens serão automaticamente classificadas como tal. O caso acima revela dois entendimentos desse aprendizado. O primeiro trata do fato de que esse aprendizado começa com resultados inadequados ou aquém do originalmente esperado. O segundo, este indispensável para a compreensão sobre o alcance e as limitações da aplicação desse aprendizado, refere-se à constatação de que esses algoritmos não estão aprendendo no sentido cognitivo, estão apenas aprimorando sua precisão para detectar padrões a partir do aumento de dados disponíveis.

A aprendizagem não supervisionada, por sua vez, envolve o processamento de dados não rotulados, ou seja, a máquina identifica, classifica e agrupa os dados de entrada (*input*) sem interação de agente externo. Com a mesma autonomia, partindo de inferências construídas pela própria máquina, baseada na conjugação de dados analisados¹⁶ e padrões identificados, dar-se-á o *output*. Em função da complexidade desse aprendizado extraído de observação, identificação de padrões e inferências, bem como das diversas camadas de relações estabelecidas por esse processamento, esse algoritmo recebe o nome de *deep learning* (aprendizagem profunda). Bom exemplo da aplicação dessa técnica reside em análises que dependem da detecção de padrões ocultos, tanto em cenário cuja descoberta de padrões seja o objetivo, quanto em que seja o meio para alcançar a determinada finalidade. Dentre as diversas utilizações para as técnicas de aprendizagem profunda, destaca-se o uso para identificação de faces, para o diagnóstico de doenças, para a criação de drones, entre outras.

Por último, sobreluz a técnica de aprendizagem por reforço,¹⁷ na qual os algoritmos não aprendem a partir de um conjunto de dados de treinamento preestabelecido, mas a partir dos dados resultantes de um ambiente de aprendizagem. Diferentemente das técnicas que alimentam o algoritmo com *feedback* instrutivo por meio de grandes quantidades de dados não rotulados ou do fornecimento de exemplos de saídas corretas para entradas fornecidas, o aprendizado por reforço depende de *feedback* avaliativo de quanto uma saída é correta. Para a compressão dessa técnica, frequentemente é utilizado o problema do *multi-armed bandit* (bandido multiarmado), que recebe esse nome porque se baseia em máquinas de caça-níqueis, também chamadas de *one-armed bandit* (bandido de um braço). Na analogia, o agente aprendiz se depara com diversas máquinas caça-níqueis (que

¹⁶ VALLE, Vanice Lírio do. Inteligência artificial incorporada à Administração Pública: mitos e desafios teóricos. *A&C – Revista de Direito Administrativo & Constitucional*, Belo Horizonte, ano 20, n. 81, p. 179-200, jul./set. 2020.

¹⁷ AMARILES, D. R. Algorithmic Decision Systems Automation and Machine Learning in the Public Administration. In: *The Cambridge Handbook of The Law of Algorithms*. Cambridge: Cambridge University Press, 2021. p. 307.

são programadas para dar recompensas a cada “x” tentativas) sem ter informações pretéritas sobre o uso dos equipamentos e, com base na tentativa e erro, precisa maximizar suas recompensas. Em suma, essa técnica de aprendizado baseia-se no treinamento para escolher ações com o único objetivo de maximizar as recompensas.

Na prática, esses sistemas de aprendizado funcionam estruturados nas chamadas redes neurais artificiais. A título de ilustração, a figura abaixo foi adaptada do estudo coordenado pelo Dr. Glenn G. Gardner.¹⁸ Nesse estudo, o objetivo era analisar imagens do globo ocular dos pacientes para detectar a existência de retinopatia, doença que causa lesões inflamatórias na parte detrás do olho (retina).

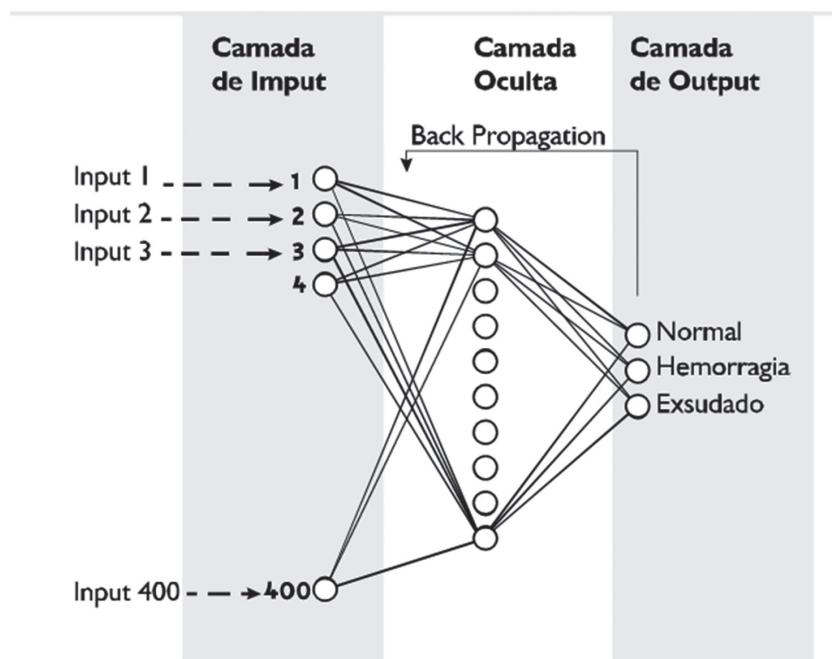


Ilustração de rede neural artificial com back propagation.

No caso em tela, o sistema de aprendizagem supervisionada recebeu *input* de conjunto de dados (imagens) para treinar o aprendizado sobre os aspectos relativos ao olho saudável e doente. A cada *input*, o sistema realizava cálculos matemáticos para encontrar padrões e, a cada erro ou padrão determinado, o sistema propagava esse resultado para que o novo processamento considerasse essa memória.

O exemplo acima serve para revelar dois importantes entendimentos sobre a decisão algorítmica: (i) em cada processamento, o *output* é resultante de variáveis

¹⁸ GARDNER, G. G.; KEATING, D.; WILLIAMSON, T. H.; ELLIOTT, A. T. Automatic detection of diabetic retinopathy using an artificial neural network: a screening tool. *British Journal of Ophthalmology*, v. 80, n. 11, p. 940-944, nov. 1996.

diferentes, ou seja, dos parâmetros redefinidos a partir da última propagação de aprendizado (*back propagation*); (ii) ter acesso ao código-fonte do algoritmo não significa ter acesso à estrutura de tomada de decisão – pois esta é dinâmica, em constante aprimoramento por meio da propagação de experiências pretéritas –, e sim à regra de aprendizagem.

Por último, ressalta-se que o exemplo trabalhado acima seria potencialmente mais complexo e profundo se tratasse de aprendizado não supervisionado (*deep learning*), pois suas redes neurais artificiais são exponencialmente mais complexas, assim como seu processamento; portanto, qualquer tentativa de análise dessas estruturas de decisão torna-se virtualmente inviável ou infrutífera.

Diante de todo o exposto, percebe-se que ambos os métodos encontram similaridades com mecanismos de aprendizagem dos seres humanos: um bebê, nos primeiros meses de vida, aprende por meio do *deep learning* (identificação de padrões e inferências sem instruções recebidas de agente externo) que, ao chorar, receberá atendimento (companhia, alimento, conforto, etc.), assim como, por meio de instruções recebidas de um agente externo, aprende, nos primeiros anos de vida, a pedir para ir ao banheiro, gerenciando suas necessidades fisiológicas, ou a amarrar o tênis, ou a utilizar talheres, etc. Assim sendo, o desafio que se apresenta na evolução da IA é o aprendizado que possibilite à leitura de contexto,¹⁹ ou seja, que incorpore elementos intuitivos de percepção de sistêmica de elementos que componham determinado cenário, a exemplo da distinção entre lágrimas de alegria e de tristeza, entre o sorriso alegre, malicioso, constrangido, entre outras percepções que dependam de cada caso em concreto. Tendo em vista o atual avanço no reconhecimento de fala e de voz, essa nova fronteira, mesmo que distante, revela-se cada vez mais tangível.

2 O funcionamento da tomada de decisão e o dilema dos algoritmos decisórios

Em ambos os métodos de aprendizagem, é perceptível que a IA aprenderá a partir de um conjunto de dados (*dataset*) provenientes do mundo físico, sejam eles inseridos e rotulados por um operador externo, sejam captados e classificados de forma autônoma pelo algoritmo. Dentre os diversos cenários de risco que se apresentam, destacam-se os efeitos provenientes da utilização de conjuntos de dados viciados e do uso de *dataset* não inclusivo, bem como os efeitos da opacidade.

¹⁹ VALLE, Vanice Lírio do. Inteligência artificial incorporada à Administração Pública: mitos e desafios teóricos. *A&C – Revista de Direito Administrativo & Constitucional*, Belo Horizonte, ano 20, n. 81, p. 179-200, jul./set. 2020.

Conjunto de dados viciados: em 2016, a Microsoft lançou um *bot* chamado Tay com o objetivo de gerar seu aprendizado a partir de conversas com pessoas usuárias da rede social Twitter. Em menos de 24 horas, o *bot* alterou seu comportamento educado e cortês para agressivo, iniciando postagens nazistas e adotando postura racista e machista.²⁰

Dataset não inclusivo: Joy Buolamwini²¹ – enquanto graduanda de ciências da computação – desenvolveu um robô social com o objetivo de interagir por meio de uma brincadeira infantil comum na Inglaterra, em que o participante cobre o rosto e, ao descobri-lo, diz “*peekaboo*”. Para realização dessa atividade, o *software* desenvolvido precisaria identificar o rosto descoberto. Embora o projeto tenha obtido êxito em experiências com outras pessoas, não funcionava quando Joy experimentava. A graduanda percebeu então que o *software* não a reconhecia porque ela era negra, e o *software* aberto utilizado como base para o reconhecimento facial não identificava rostos de pessoas negras. Para apresentar seu projeto, Joy precisava utilizar uma máscara branca ou pedir para alguma pessoa de pele clara testar o experimento. Em apresentação ao TED Talk, Joy contou que, mais tarde, cursando seu Ph.D no *Massachusetts Institute of Technology*, participou de uma competição em Hong Kong em que um grupo participante utilizava tecnologia de reconhecimento facial, porém o rosto de Joy novamente não foi identificado. Após isso, ela constatou que o *software* utilizado pelo grupo era o mesmo que ela utilizara no passado.

Opacidade: em 2013, no estado americano de Wisconsin, Eric Loomis foi preso em flagrante e conduzido ao juiz para determinar se responderia ao processo em liberdade (caso considerasse baixo o risco de reincidência) ou sob custódia do Estado. O juiz negou sua liberdade provisória fundamentando a decisão com base no *output* do *software* COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*). A defesa pleiteou a devida motivação objetiva para essa decisão, a qual foi negada em primeira instância; ato contínuo, recorreram à Suprema Corte de Wisconsin, que decidiu negar o acesso às informações, justificando esse juízo pelo fato de que o *software* seria protegido pelo sigilo industrial (*software* proprietário). Em último esforço, a defesa à Suprema Corte do país – tendo a prerrogativa legal de escolher os casos que julga – decidiu não tratar da causa em questão por tratar-se de assunto novo, de pouca familiaridade para justificar o juízo da corte. Após o caso, constatou-se que o *software* funcionava a partir de um algoritmo secreto e que nem os juízes que o utilizavam tinham ciência de como o

²⁰ MOREIRA, Isabela. A Microsoft criou uma robô que interage nas redes sociais – e ela virou nazista. *Revista Galileu*, 24. mar. 2016. Disponível em: <https://revistagalileu.globo.com/blogs/buzz/noticia/2016/03/microsoft-criou-uma-robo-que-interage-nas-redes-sociais-e-ela-virou-nazista.html>. Acesso em: 18 ago. 2021.

²¹ BUOLAMWINI, Joy. How I'm fighting bias in algorithms. *TED*, nov. 2016. Disponível em: https://www.ted.com/talks/joy_buolamwini_how_i_m_fighting_bias_in_algorithms?language=pt-BR. Acesso em: 18 ago. 2021.

software funcionava ou de como gerava os resultados propostos. A partir disso, foram realizadas auditorias nos resultados do COMPAS pela ONG ProPublica²² revelando que o *software* era enviesado contra pessoas negras de tal forma que essas pessoas teriam duas vezes mais chances que pessoas brancas de serem consideradas de alto risco. Embora exista a observação de que o algoritmo não possui o critério racial para a tomada de decisão, acredita-se que – assim como outros sistemas de segurança utilizados nos EUA, como PredPol,²³ da cidade Chicago, entre outros – um *input* como o CEP pode ser utilizado como indicador de grupo étnico, sobretudo em países segregados, para os EUA.

O levantamento de apenas esses três itens (dados viciados, dados não inclusivos e a opacidade inerente ao processo decisório da IA) revela desafios de grande escala, a partir dos quais se discorrerá.

3 A compreensão dos vieses decisórios algorítmicos

Compreendidos os elementos técnicos de automação da aprendizagem e vistas as formas de funcionamento da tomada de decisão algorítmica, inegável a percepção de que existe uma interação entre o homem e a inteligência artificial, a qual merece atenção no presente momento.

A relação entre o homem e a inteligência artificial tem sido tema de muitos debates e controvérsias no que se refere aos potenciais de evolução tecnológica e suas interações com a sociedade.

Em entrevista, o ministro Luiz Fux, do Supremo Tribunal Federal, apontou três reflexões sobre o tema aos quais – segundo ele – os estudiosos devem se ater: “O alcance e o escopo da aplicação da Inteligência Artificial no mundo jurídico; desafios e problemas éticos relacionados à regulação dessas novas tecnologias; e exemplos de sucesso da aplicação dessa tecnologia no Judiciário brasileiro”.²⁴

Nesse sentido, o foco do presente artigo está nos aspectos relacionados ao dilema ético dessa relação entre homem e inteligência artificial. Por isso, surgem indagações no sentido de: poderá o homem ser substituído pela inteligência artificial? A máquina pode potencializar vieses da sociedade humana ou ela será imparcial e capaz de corrigi-los? Quais medidas podem ser adotadas para desviesar os vieses humanos e aqueles identificados por meio das máquinas?

²² ANGWIN, Julia; LARSON, Jeff; MATTU, Surya; KIRCHNER, Lauren. Machine bias. *ProPublica*, 23 maio 2016. Disponível em: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. Acesso em: 05 ago. 2021.

²³ O'NEIL, C. *Algoritmos de destruição em massa: como o big data aumenta a desigualdade e ameaça a democracia*. Santo André: Rua do Sabão, 2020. p. 97.

²⁴ FUX mostra benefícios e questionamentos da inteligência artificial no Direito. *Consultor Jurídico*, 14 mar. 2019. Disponível em: <https://consultor-juridico.jusbrasil.com.br/noticias/685433315/fux-mostra-beneficios-e-questionamentos-da-inteligencia-artificial-no-direito#>. Acesso em: 20 ago. 2021.

Diante disso, para responder aos questionamentos e, por conseguinte, buscar apresentar uma proposta de solução a esse dilema ético da decisão algorítmica, o estudo merece se ater, neste primeiro ponto, à compreensão dos vieses existentes na sociedade.

3.1 Da compreensão dos vieses humanos e da inteligência artificial

A definição e compreensão do que seria um viés, antes de mais nada, afigura-se como de extrema relevância para a fundamentação que se pretende expor a partir de agora sobre o tema.

Por conta disso, quando chamada a atenção sobre o que seriam os vieses humanos ou de máquinas (inteligência artificial), a interpretação que se busca está no conceito de que os vieses seriam uma visão tendenciosa ou propensão desvirtuada sobre determinado tema ou assunto que se pretende discutir.

A despeito da propensão desvirtuada dos vieses, a qual será objeto de análise do presente artigo, não desconsideramos que existem vieses benéficos e que podem apresentar resultantes objetivas boas, como visto nos exemplos já mencionados sobre direcionamento de conteúdos mais consumidos por usuários de certo aplicativo.

Compreendida essa definição inicial sobre os vieses, portanto, tem-se por certo que a sociedade – e aqui, por óbvio, está inserido o homem – possui interpretação já predefinida quanto a determinados assuntos ou temas, a qual pode ou não caracterizar um pré-conceito enviesado nas gerações humanas. E essa percepção sobre uma sociedade enviesada não passou despercebida dos ensinamentos do professor Juarez Freitas,²⁵ para quem, “desde sempre, os vieses condicionam, moldam e impelem a tomada de decisão, na história das civilizações. Quando automatismos fortemente obnublam a esfera decisória humana, a liberdade, a equidade e a dignidade entram em ocaso”.

Desse modo, o conhecimento sobre o enviesamento humano e as dificuldades dele decorrentes, haja vista a necessidade de se obterem decisões éticas imparciais, não podem deixar de ser considerados quando se está diante de uma decisão algorítmica, principalmente pelo fato de que a própria máquina pode potencializar o enviesamento.

Melhor explicando, sendo de conhecimento que a sociedade possui vieses já concebidos, a adoção de um sistema de inteligência artificial, cuja finalidade será processar um conjunto de dados (“*dataset*”, resultantes das relações enviesadas

²⁵ FREITAS, Juarez; FREITAS, Thomas Bellini. *Direito e inteligência Artificial*. Belo Horizonte: Fórum, 2020. p. 93.

dessa sociedade) – sejam estes inseridos ou rotulados por um agente externo, sejam advindos do processamento autônomo (*deep learning*) –, tem-se por certo a necessidade de se ponderar sobre o risco da potencialização do enviesamento nas resultantes desse processamento (*output*).

Os riscos dessa situação tendem a revelar e intensificar o enviesamento humano ao se processarem as informações inseridas no sistema de inteligência artificial. Em função disso, Fausto de Sanctis²⁶ chama atenção para o ponto ao afirmar que as funções serão modificadas com a inteligência artificial, oportunidade em que haverá dificuldades nessa relação entre o enviesamento humano e a inteligência artificial: “Katie Brigham revela ser fácil imaginar que existe uma maneira melhor, que um dia encontraremos uma ferramenta que possa tomar decisões neutras e desapaixonadas sobre investigações e punição, entretanto, não se pode desconsiderar que humanos são imperfeitos porquanto sujeitos a preconceitos e estereótipos e, quando estes entram em jogo no sistema de justiça criminal, as pessoas mais desfavorecidas acabam sofrendo”.

Exemplo dessa preocupação apresentada pelo autor, acima citada, encontra-se no *case* envolvendo o COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*), o qual foi criado com a intenção de tornar decisões judiciais menos subjetivas e afastadas de erros humanos, preconceitos ou racismo (vieses dos humanos).

A despeito da boa intenção do sistema, quando do processamento das informações obtidas pelo sistema por meio da inteligência artificial, este passou a atribuir a minorias étnicas pontuação superior, o que demonstrou nítida potencialização do enviesamento humano pela inteligência artificial, como pode ser identificado em notícia veiculada em 2016 – três anos após o caso Eric Loomis, acima mencionado – por meio do *site* da *BBC News*:²⁷ “Quando analisamos um acusado negro e outro branco com a mesma idade, sexo e ficha criminal – e levando em conta que depois de serem avaliados os dois cometeram quatro, dois ou nenhum crime –, o negro tem 45% mais chances do que o branco de receber uma pontuação alta”, afirma Julia Angwin.

Assim, compreende-se que o enviesamento humano existe e que, quando sua aparição ocorre no sistema de inteligência artificial, aquele pode ser potencializado de forma a evidenciar os vieses humanos.

Por sua vez, a análise do *dataset* pode ser processada de tal forma que, mesmo não sendo o interesse humano, se mostra capaz de interpretar determinado

²⁶ SANCTIS, Fausto de. *Inteligência Artificial e Direito*. São Paulo: Almedina, 2020. p. 115.

²⁷ MAYBIN, Simon. Sistema de algoritmo que determina pena de condenados cria polêmica nos EUA. *BBC News Brasil*, 31 out. 2016. Disponível em: <https://www.bbc.com/portuguese/brasil-37677421>. Acesso em: 20 ago. 2021.

dado de forma desvirtuada, o que também qualifica o enviesamento da própria inteligência artificial.

Com base nesses enviesamentos (humano e da inteligência artificial), os quais podem, como apontado, criar violações aos aspectos éticos em decisões algorítmicas, temos que a solução do dilema ético da decisão algorítmica consiste na busca de um antídoto para resolver os vieses e alcançar a ética.

3.2 A busca da decisão algorítmica ética por meio do desenviesamento humano e da inteligência artificial

Se existentes o enviesamento humano e sua possibilidade de potencialização, mediante identificação pela própria inteligência artificial ao processar os dados, a busca por uma decisão algorítmica ética consiste na identificação de um antídoto capaz de afastar os problemas enfrentados no tópico anterior.

Ao que parece, referido antídoto consiste no chamado desenviesamento dos vieses humanos e da inteligência artificial.

De acordo com os ensinamentos de Juarez Freitas,²⁸ “o desenviesamento (ou a reprogramação intencional das predisposições discriminatórias e injustas) é uma agenda mandatária”. Ou seja, “automatismos têm que ser criticamente reciclados para vetar os condicionamentos nefastos – ostensivos ou velados – que minam o livre-arbítrio e o melhor da humanidade”.

Ainda sobre esse desenviesamento e fazendo uma interpretação mais restrita ao âmbito das decisões proferidas por juízes, novamente os ensinamentos de Fausto de Sanctis²⁹ auxiliam nas dificuldades e propostas de soluções para o tema: “A Inteligência Artificial não significa apenas uma palavra da moda. Ela interfere na vida de todos com capacidade de transformar vários procedimentos antes impensáveis. O uso da tecnologia, além de buscar a celeridade dos julgamentos, visa remover o erro humano, retirando-o da equação. Em outras palavras, a experiência humana alimentada nos sistemas poderia nos oferecer decisões judiciais objetivas, ou seja, sem vieses porque fruto de nossos preconceitos, dramas e angústias. [...] A Inteligência Artificial seria, então o remédio perfeito para retirar o subjetivismo judicial e, com isso, as falhas humanas. Em verdade, em muitas hipóteses, pode melhorar o sistema, mas não em todas”.

A relação entre inteligência artificial e humanos, portanto, tem muitos motivos para ser vista com um potencial de melhoria, mesmo diante dos diversos desafios.

²⁸ FREITAS, Juarez; FREITAS, Thomas Bellini. *Direito e Inteligência Artificial*. Belo Horizonte: Fórum, 2020. p. 93.

²⁹ SANCTIS, Fausto de. *Inteligência Artificial e Direito*. São Paulo: Almedina, 2020. p. 116.

A experiência passada, como sempre ocorre nas evoluções sociais, serve de aprendizado para melhorias para o futuro. Por isso, não se pode ter a compreensão de que a integração entre humano e inteligência artificial é algo ruim. Muito pelo contrário, pois, assim como não se pode atrelar toda a decisão com fundamentos apresentados somente pelo sistema de inteligência artificial, igualmente não pode o humano afastar todas as informações que foram apontadas objetivamente por meio do processamento das informações, justificando a busca por um equilíbrio nessa relação.

Como afirma Fausto de Sanctis, o “objetivo é ajudar os seres humanos e as máquinas a colaborarem com o sucesso, contemplando uns aos outros com objetividade técnica e experiência humana”.³⁰

E é a essa integração entre a inteligência artificial e o homem, a qual exige a interpretação humana em relação aos dados e informações processados pelo sistema de inteligência artificial, que o professor Juarez Freitas afirmou no sentido de que há um “duplo movimento”.³¹ Neste, esclarece o autor, que deverá ser preconizado que, por um lado, haverá necessidade de se buscar o desviesamento da inteligência artificial, ao passo que, de outro vértice, a própria inteligência artificial deve processar os dados e informações para afastar os vieses humanos e apresentar resultados desviesados.

Com base nessa contraposição, ainda, é que o autor afirma que o “ponto de inflexão está em dotar a IA de imparcialidade suficiente para analisar criticamente os dados e, ao mesmo tempo, submetê-los a um programa desviesante (atento ao protocolo recomendado)”.³²

Desse modo, quando se fala na busca por uma decisão algorítmica ética, deve-se ter em mente que esta consistirá em um desviesamento humano e da inteligência artificial, cuja finalidade será alcançar decisões mais justas e éticas, mediante respeito e observância de preceitos constitucionais e legais.

Em síntese, considerando que o desviesamento parece ser meio adequado para o afastamento de vieses humanos e da inteligência artificial, faz-se necessária a compreensão das normas brasileiras que asseguram e buscam efetivar essa conduta desviesante.

³⁰ SANCTIS, Fausto de. *Inteligência Artificial e Direito*. São Paulo: Almedina, 2020. p. 116.

³¹ FREITAS, Juarez; FREITAS, Thomas Bellini. *Direito e inteligência Artificial*. Belo Horizonte: Fórum, 2020. p. 96.

³² FREITAS, Juarez; FREITAS, Thomas Bellini. *Direito e inteligência Artificial*. Belo Horizonte: Fórum, 2020. p. 99.

4 A compreensão atual do sistema brasileiro e seus aspectos desviesantes para superar o dilema ético da decisão algorítmica

Na medida em que a controvérsia acerca dos vieses e dos aspectos referentes ao desviesamento foi delimitada anteriormente, tem-se que ter em mente que a compreensão do sistema brasileiro passa pelos desafios impostos à administração pública em assegurar decisões algorítmicas éticas. Para tanto, parte-se do pressuposto de que a base primária para análise dos parâmetros a serem observados pela administração pública é a Constituição Federal, a qual assegura à sociedade o direito fundamental à boa administração, de maneira a torná-la mais eficiente e eficaz.

Nesse sentido, para que se tenha uma decisão algorítmica ética no âmbito da esfera pública, o direito fundamental à boa administração deve ser interpretado como obrigatório para afastar o enviesamento humano e da inteligência artificial. Sobre essa obrigatoriedade, o professor Juarez Freitas³³ é preciso ao dizer que referida situação exige um “direito à administração pública imparcial e o mais desviesada possível, isto é, aquela que, filtrando os desvios cognitivos, não pratica nem estimula discriminação negativa de qualquer natureza e, ao mesmo tempo, promove discriminações inversas ou positivas (reduzoras das desigualdades iníquas)”.

Somada a essa perspectiva constitucional, a Lei de Introdução às Normas de Direito Brasileiro (LINDB) foi reforçada no que se refere às normas de direito público, haja vista a publicação da Lei Federal nº 13.655, de 25 de abril de 2018.³⁴

Quanto ao referido reforço, por sua vez, merecem destaque os termos do artigo 30 da LINDB, que trouxe a previsão de que “[a]s autoridades públicas devem atuar para aumentar a segurança jurídica na aplicação das normas, inclusive por meio de regulamentos, súmulas administrativas e respostas a consultas”.

Esse aumento da segurança jurídica na administração pública com base no advento do artigo 30 da LINDB foi muito bem explicado em artigo do professor Egon Bockmann Moreira com Paula Pessoa Pereira,³⁵ os quais afirmam que referido

³³ FREITAS, Juarez. As políticas públicas e o direito fundamental à boa administração. *Nomos: Revista do Programa de Pós-Graduação em Direito da UFC*, Fortaleza, v. 35, n. 1, p. 195-217, jan./jun. 2015. p. 198-199.

³⁴ GABARDO, Emerson; SOUZA, Pablo Ademir de. O consequencialismo e a LINDB: a cientificidade das previsões quanto às consequências práticas das decisões. *A&C – Revista de Direito Administrativo & Constitucional*, Belo Horizonte, ano 20, n. 81, p. 97-124, jul./set. 2020; VITORELLI, Edilson. A Lei de Introdução às Normas do Direito Brasileiro e a ampliação dos parâmetros de controle dos atos administrativos: um novo paradigma. *A&C – Revista de Direito Administrativo & Constitucional*, Belo Horizonte, ano 19, n. 78, p. 195-219, out./dez. 2019.

³⁵ MOREIRA, Egon Bockmann; PEREIRA, Paula Pessoa. Art. 30 da LINDB - O dever público de incrementar a segurança jurídica. *Revista de Direito Administrativo*, Rio de Janeiro, Edição Especial: Direito público na Lei de Introdução às Normas do Direito Brasileiro – LINDB (Lei nº 13.655/2018), p. 243-274, 2018.

dispositivo possui dois desdobramentos – “efetivo respeito” e “estabilização institucional” – em relação às decisões.

No que se refere ao “efetivo respeito” às decisões, visa-se afastar todas as interferências externas ao julgar casos semelhantes, o que justifica um tratamento igualitário entre todos; quanto à “estabilização institucional” das decisões, somente haverá estabilidade se houver transparência e atuação legítima das instituições, a lembrar que a confiança será afastada nos casos em que obscuridades afetarem os termos decisórios.

Diante dessas considerações, o que se propõe para a busca de uma decisão algorítmica ética é a interpretação sistêmica da Constituição Federal com os preceitos estabelecidos na LINDB, em especial a regra especificada no artigo 30.

Para afastar os vieses humanos e da inteligência artificial, o desviosamento deve ser tratado como assunto sério e visto como direito fundamental da sociedade, sendo certo que a segurança jurídica, ao se proferirem decisões algorítmicas, ocorrerá só quando os procedimentos e as regras forem transparentes, imparciais e éticas.

De modo a confirmar essa proposição, ao que parece, o Brasil vem adotando, no âmbito da administração pública, diretrizes para alcançar o desviosamento nas decisões algorítmicas a serem proferidas, como é o caso, por exemplo, dos preceitos expressos na Resolução nº 332, de 21 de agosto de 2020, do Conselho Nacional de Justiça.

Nessa resolução, o desenvolvimento e a implantação da inteligência artificial devem ser compatíveis com os direitos fundamentais, inclusive sendo possível colher das afirmações ali previstas que a tomada de decisão deverá respeitar critérios éticos e de transparência.

Aliás, a seriedade no tratamento do tema parece estar confirmada quando do artigo 7º, §3º, da resolução, que diz que “[a] impossibilidade de eliminação do viés discriminatório do modelo de Inteligência Artificial implicará na descontinuidade de sua utilização, com o conseqüente registro de seu projeto e as razões que levaram a tal decisão”.

Outrossim, tramita no Senado Federal o Projeto de Lei nº 5.691, de 2019,³⁶ o qual pretende instituir a Política Nacional de Inteligência Artificial e tem como objetivo estimular a formação de um ambiente favorável ao desenvolvimento das tecnologias em inteligência artificial.

A preocupação com a interpretação sistêmica aqui proposta, por sua vez, é a mesma, ao que parece, adotada para a instituição da Política Nacional, principalmente

³⁶ BRASIL. Senado Federal. *Projeto de Lei n. 5.691, de 2019*. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/139586>. Acesso em: 20 ago. 2021.

pelo fato de essa estabelecer, dentre os seus princípios, os da transparência, segurança e confiabilidade (atual artigo 2º, inciso IV, do projeto de lei).

Assim, o dilema ético das decisões algorítmicas passa por uma interpretação sistêmica do ordenamento jurídico pátrio, de modo que somente haverá segurança e transparência naquelas se e quando existentes normas que visem ao desviosamento humano e da inteligência artificial, a fim de assegurar o direito fundamental a uma boa administração pública.

5 Conclusão

O atual estágio de evolução da inteligência artificial, como visto durante toda a exposição, fez com que esta passasse a ser integrada como suporte à tomada de decisão nas diversas esferas e dimensões da vida em sociedade.

A despeito desse suporte, diversos são os desafios dessa integração, porque, seja diante da aprendizagem supervisionada, seja por meio da aprendizagem não supervisionada (*deep learning*), verificou-se que os vieses tendem a ser potencializados após o processamento dos dados (*output*).

Referidos vieses, os quais podem ser identificados por meio do *dataset*, oriundo do enviesamento humano, surgem em razão dos vícios, pré-conceitos e costumes institucionalizados, motivo pelo qual os diversos desafios estão inseridos na solução desses vieses.

O antídoto para esses e a finalidade de alcançar decisões algorítmicas éticas, transparentes e imparciais podem ser vistos por meio do desviosamento.

Em função disso, o duplo movimento pode ser adotado como forma de afastar os vieses que serão potencializados como resultantes do processamento de dados (*output*). No primeiro movimento, o desviosamento³⁷ deve ocorrer na programação, no curso da aprendizagem e em correções posteriores; num segundo movimento, preceitua-se o uso da IA como instrumento para detectar e retificar enviesamentos insidiosos.

O equilíbrio entre homem e máquina, mediante adoção desse duplo movimento, apresenta-se como meio para possível solução dos diversos desafios que se apresentam.

Tendo essa percepção em mente, o Brasil está caminhando para adotar esse duplo movimento. A interpretação sistêmica entre Constituição Federal e artigo 30 da LINDB reforça o posicionamento de que a solução dos diversos desafios passa pela compreensão de que é direito fundamental da sociedade exigir a boa administração pública, bem como ter segurança quanto ao processamento de dados

³⁷ FREITAS, Juarez; FREITAS, Thomas Bellini. *Direito e inteligência Artificial*. Belo Horizonte: Fórum, 2020. p. 96.

e ao uso da IA nas suas diversas aplicações, assim como exigir transparência, que, em se tratando de algoritmos, significa acessibilidade e compreensibilidade.³⁸

Assim, o dilema ético da decisão algorítmica está em pauta, e a solução para tal consiste em tratar o assunto como um direito fundamental da sociedade de se ter uma boa administração pública, que passe confiança, transparência e segurança em suas decisões a serem proferidas.

Referências

AMARILES, D. R. Algorithmic Decision Systems Automation and Machine Learning in the Public Administration. In: *The Cambridge Handbook of The Law of Algorithms*. Cambridge: Cambridge University Press, 2021.

ANGWIN, Julia; LARSON, Jeff; MATTU, Surya; KIRCHNER, Lauren. Machine bias. *ProPublica*, 23 maio 2016. Disponível em: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. Acesso em: 05 ago. 2021.

ARAÚJO, Valter Shuenquener de; ZULLO, Bruno Almeida; TORRES, Maurílio. Big Data, algoritmos e inteligência artificial na Administração Pública: reflexões para a sua utilização em um ambiente democrático. *A&C – Revista de Direito Administrativo & Constitucional*, Belo Horizonte, ano 20, n. 80, p. 241-261, abr./jun. 2020.

BELLOCCHIO, Lucía; SANTIAGO, Alfonso. Estado digital de Derecho. *A&C – Revista de Direito Administrativo & Constitucional*, Belo Horizonte, ano 20, n. 80, p. 87-102, abr./jun. 2020.

BRASIL. Senado Federal. *Projeto de Lei n. 5.691, de 2019*. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/139586>. Acesso em: 20 ago. 2021.

BRAVO, Álvaro Avelino Sánchez. Marco Europeo para una inteligencia artificial basada en las personas. *International Journal of Digital Law*, Curitiba, v. 1, n. 1, p. 65-77, jan./abr. 2020.

BUOLAMWINI, Joy. How I'm fighting bias in algorithms. *TED*, nov. 2016. Disponível em: https://www.ted.com/talks/joy_buolamwini_how_i_m_fighting_bias_in_algorithms?language=pt-BR. Acesso em: 18 ago. 2021.

CORVALÁN, Juan Gustavo. Inteligencia Artificial GPT-3, Pretoria y Oráculos Algorítmicos en el Derecho. *International Journal of Digital Law*, Curitiba, v. 1, n. 1, p. 11-52, jan./abr. 2020.

FERRARI, I. *Justiça digital* [livro eletrônico]. 2. ed. São Paulo: Thomson Reuters Brasil, 2021.

FIGUEIREDO, Carla Regina Bortolaz de; CABRAL, Flávio Garcia. Inteligência artificial: machine learning na Administração Pública. *International Journal of Digital Law*, Curitiba, v. 1, n. 1, p. 79-96, jan./abr. 2020.

FORNASIER, Mateus de Oliveira. The applicability of the Internet of Things (IoT) between fundamental rights to health and to privacy. *Revista de Investigações Constitucionais*, Curitiba, vol. 6, n. 2, p. 297-321, maio/ago. 2019.

³⁸ FERRARI, I. *Justiça digital* [livro eletrônico]. 2. ed. São Paulo: Thomson Reuters Brasil, 2021. RB-7.3.

FREITAS, Juarez. As políticas públicas e o direito fundamental à boa administração. *Nomos: Revista do Programa de Pós-Graduação em Direito da UFC*, Fortaleza, v. 35, n. 1, p. 195-217, jan./jun. 2015.

FREITAS, Juarez; FREITAS, Thomas Bellini. *Direito e inteligência artificial*. Belo Horizonte: Fórum, 2020.

FUX mostra benefícios e questionamentos da inteligência artificial no Direito. *Consultor Jurídico*, 14 mar. 2019. Disponível em: <https://consultor-juridico.jusbrasil.com.br/noticias/685433315/fux-mostra-beneficios-e-questionamentos-da-inteligencia-artificial-no-direito#>. Acesso em: 20 ago. 2021.

GABARDO, Emerson; SOUZA, Pablo Ademir de. O consequencialismo e a LINDB: a cientificidade das previsões quanto às consequências práticas das decisões. *A&C – Revista de Direito Administrativo & Constitucional*, Belo Horizonte, ano 20, n. 81, p. 97-124, jul./set. 2020.

GALLO APONTE, W. I.; LÓPEZ VALLE, V. C.; FÁCIO, R. N. La utilización de inteligencia artificial en la actividad regulatoria: una propuesta en favor del desarrollo nacional sostenible. *Veredas do Direito*, Belo Horizonte, v. 17, n. 39, p. 123-146, sep./dic. 2020.

GARDNER, G. G.; KEATING, D.; WILLIAMSON, T. H.; ELLIOTT, A. T. Automatic detection of diabetic retinopathy using an artificial neural network: a screening tool. *British Journal of Ophthalmology*, v. 80, n. 11, p. 940-944, nov. 1996.

MAYBIN, Simon. Sistema de algoritmo que determina pena de condenados cria polêmica nos EUA. *BBC News Brasil*, 31 out. 2016. Disponível em: <https://www.bbc.com/portuguese/brasil-37677421>. Acesso em: 20 ago. 2021.

MOREIRA, Egon Bockmann; PEREIRA, Paula Pessoa. Art. 30 da LINDB - O dever público de incrementar a segurança jurídica. *Revista de Direito Administrativo*, Rio de Janeiro, Edição Especial: Direito público na Lei de Introdução às Normas do Direito Brasileiro – LINDB (Lei nº 13.655/2018), p. 243-274, 2018.

MOREIRA, Isabela. A Microsoft criou uma robô que interage nas redes sociais – e ela virou nazista. *Revista Galileu*, 24. mar. 2016. Disponível em: <https://revistagalileu.globo.com/blogs/buzz/noticia/2016/03/microsoft-criou-uma-robo-que-interage-nas-redes-sociais-e-ela-virou-nazista.html>. Acesso em: 18 ago. 2021.

NOHARA, Irene Patrícia; COLOMBO, Bruna Armonas. Tecnologias cívicas na interface entre direito e inteligência artificial: Operação Serenata de Amor para góstosuras ou travessuras? *A&C – Revista de Direito Administrativo & Constitucional*, Belo Horizonte, ano 19, n. 76, p. 83-103, abr./jun. 2019.

O'NEIL, C. *Algoritmos de destruição em massa: como o big data aumenta a desigualdade e ameaça a democracia*. Santo André: Rua do Sabão, 2020.

SANCTIS, Fausto de. *Inteligência Artificial e Direito*. São Paulo: Almedina, 2020.

SILVA, L. A. D. Uso de técnicas de inteligência artificial para subsidiar ações de controle. *Revista do TCU*, Brasília, n. 137, p. 124-129, set./dez. 2016.

STRINGHINI, Antonella. Administración Pública Inteligente: novedades al ecosistema normativo digital de la República Argentina. *Revista Eurolatinoamericana de Derecho Administrativo*, Santa Fe, vol. 5, n. 2, p. 199-215, jul./dic. 2018.

STRINGHINI, Antonella. Asistencia virtual automatizada e inclusiva para optimizar la relación de la ciudadanía con la Administración Pública. *International Journal of Digital Law*, Curitiba, v. 1, n. 1, p. 117-128, jan./abr. 2020.

VALENTE, J. Órgãos públicos usam inteligência artificial para combater corrupção. *Agência Brasil*, 3 ago. 2018. Disponível em: <https://agenciabrasil.ebc.com.br/geral/noticia/2018-08/orgaos-publicos-usam-inteligencia-artificial-para-combater-corrupcao>. Acesso em: 1º mar. 2020.

VALLE, Vanice Lírio do. Inteligência artificial incorporada à Administração Pública: mitos e desafios teóricos. *A&C – Revista de Direito Administrativo & Constitucional*, Belo Horizonte, ano 20, n. 81, p. 179-200, jul./set. 2020.

VALLE, Vivian Cristina Lima López; GALLO, William Ivan. Inteligência artificial e capacidades regulatórias do Estado no ambiente da administração pública digital. *A&C – Revista de Direito Administrativo & Constitucional*, Belo Horizonte, ano 20, n. 82, p. 67-86, out./dez. 2020.

VALLE, Vivian Cristina Lima López; MARTINELLI, Gustavo; BARBOSA FILHO, Claudio Roberto. Participação no processo legislativo como direito fundamental: um olhar sobre as ferramentas de e-democracy no Congresso Nacional Brasileiro. *Revista Eurolatinoamericana de Derecho Administrativo*, Santa Fe, vol. 7, n. 1, p. 143-166, ene./jun. 2020.

VIANA, Ana Cristina Aguilar. Transformação digital na administração pública: do governo eletrônico ao governo digital. *Revista Eurolatinoamericana de Derecho Administrativo*, Santa Fe, vol. 8, n. 1, p. 115-136, ene./jun. 2021.

VITORELLI, Edilson. A Lei de Introdução às Normas do Direito Brasileiro e a ampliação dos parâmetros de controle dos atos administrativos: um novo paradigma. *A&C – Revista de Direito Administrativo & Constitucional*, Belo Horizonte, ano 19, n. 78, p. 195-219, out./dez. 2019.

Informação bibliográfica deste texto, conforme a NBR 6023:2018 da Associação Brasileira de Normas Técnicas (ABNT):

PUSCHEL, André Felipe Silva; RODRIGUES, Roberto Tassis; VALLE, Vivian Cristina Lima López. O dilema ético da decisão algorítmica na administração pública. *A&C – Revista de Direito Administrativo & Constitucional*, Belo Horizonte, ano 22, n. 90, p. 207-226, out./dez. 2022. DOI: 10.21056/aec.v22i90.1737.
