

## 16 INTELIGÊNCIA ARTIFICIAL APLICADA AO DIREITO – APARÊNCIA E ESSÊNCIA

Cláudia Toledo<sup>1</sup>

Michelle Alves<sup>2</sup>

Paulo Roberto Fontes<sup>3</sup>

**Palavras-Chave:** Inteligência Artificial; Direitos Fundamentais; Estado Democrático de Direito.

### APRESENTAÇÃO

Neste resumo são apresentados resultados parciais das investigações realizadas pelos graduandos Michelle Alves e Paulo Roberto Fontes no Grupo de Pesquisa *Inteligência Artificial Aplicada ao Direito*, cadastrado no CNPq sob a coordenação da Prof<sup>a</sup> Cláudia Toledo.<sup>4</sup>

Esse grupo de pesquisa visa ao estudo da inteligência artificial (IA), à identificação de seus benefícios e vantagens, mas em uma análise crítica de verificação de sua conformidade às determinações do direito nacional (marcadamente às normas constitucionais, em especial, aos direitos fundamentais) e às normas internacionais (especialmente os direitos humanos). A expansão da tecnologia e ampliação da aplicação da IA a diversas áreas da vida humana são processos irreversíveis, assim como o ganho em celeridade por elas gerado. Entretanto, mostra-se imprescindível não apenas o estudo para o desenvolvimento da IA, mas também para seu controle, conforme interesses e valores humanos nucleares.

Assim, buscou-se nesta investigação a delimitação do *conceito* de IA e a identificação da *forma* como ela se estrutura, objetivando, mediante a satisfação das pretensões discursivas de *inteligibilidade* e *verdade*, o alcance de um discurso cientificamente *válido* em termos de IA. A justificativa desta pesquisa reside na constatação tanto de diversas noções do que seria IA,

---

<sup>1</sup> Professora Associada da Faculdade de Direito da Universidade Federal de Juiz de Fora (UFJF); Membro permanente do Programa de Pós-Graduação Stricto Sensu em Direito e Inovação (UFJF); Doutorado em Teoria e Filosofia do Direito pela Universidade Federal de Minas Gerais (UFMG); Estágio Pós-Doutoral na Universidade Christian-Albrecht (CAU), Kiel, Alemanha; Estágio Pós-Doutoral na Universidade Federal de Santa Catarina (UFSC).

<sup>2</sup> Graduanda em Direito pela Universidade Federal de Juiz de Fora. Aluna integrante do Grupo de Pesquisa Inteligência Artificial Aplicada ao Direito.

<sup>3</sup> Graduando em Direito pela Universidade Federal de Juiz de Fora. Aluno integrante do Grupo de Pesquisa Inteligência Artificial Aplicada ao Direito.

<sup>4</sup> DIRETÓRIO de Grupos de Pesquisa. Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Disponível em [dgp.cnpq.br/dgp/espelhogrupo/8877889481097583](http://dgp.cnpq.br/dgp/espelhogrupo/8877889481097583). Acesso em: 26 out. 2021.

quanto da reduzida preocupação científico-conceitual com a elaboração de sua definição e a investigação de sua efetiva potencialidade. Ao contrário, este é um dos tópicos paradigmáticos na expressão da inventividade humana, que apresenta desde os mais criativos modelos artísticos de ficção científica até asserções científicas equivocadas sobre as capacidades da IA no seu atual estado da arte.

## **METODOLOGIA**

Embora este grupo já tenha desenvolvido pesquisa empírica para o levantamento e análise, por exemplo, da presente utilização de programas de IA pelo Poder Judiciário brasileiro, apresenta-se aqui, como exposto, exclusivamente as principais discussões e conclusões dos estudos realizados especificamente pelos dois graduandos autores deste resumo, sob a orientação da professora coordenadora do Grupo de Pesquisa.

Assim, esta investigação em particular teve como objeto e recorte temático a delimitação do próprio *conceito* de IA, bem como relevantes questões envolvidas com sua capacidade e as tarefas que desenvolve. Para tanto, realizou-se pesquisa *bibliográfica*, com utilização de método *dedutivo* em investigação *qualitativa* e abordagem *analítica* de *livros e artigos científicos*, além de *artigos de opinião* (de jornais e revistas de grande circulação) utilizados como *fontes indiretas* de estudo. As conclusões deste estudo seguem como premissas de que partiram investigações posteriores atualmente em andamento pelo Grupo de Pesquisa.

## **RESULTADOS**

A primeira meta buscada nesta pesquisa foi a identificação do próprio *conceito* de *inteligência artificial*. Isso porque o primeiro passo para a satisfação da *inteligibilidade* do discurso – uma de suas *condições de validade* (HABERMAS, 1994, p. 153) – é a *definição conceitual do objeto* sobre o qual se discute.

De forma recorrente, IA é tratada como se fosse capacidade das máquinas de “pensarem por si mesmas”, ao desempenharem tarefas anteriormente realizadas por humanos, e consideradas como exigindo a capacidade de aprender, raciocinar e resolver problemas (EDX, 2020). Alguns teóricos chegam a abordar a IA com perspectiva eminentemente apocalíptica, mencionando a chamada Super Inteligência, um sistema altamente capaz, apto a se sobrepor à habilidade humana de decidir, em virtude de sua capacidade de lidar com carga de informações muito mais elevada do que é possível ao cérebro humano assimilar (RUSSEL, 2016, *apud*

PEIXOTO; SILVA, 2019, p. 79).

Sem dúvida, a capacidade de processamento de dados, de cruzamento e combinação de informações da máquina é inequivocamente superior à capacidade humana individual. Contudo, é imprescindível ressaltar que a máquina *não* é capaz de pensar por si mesma, pelo simples fato de que ela *não é autônoma*. O programa de IA não cria seus próprios algoritmos, eles são elaborados por seu *desenvolvedor*. Conforme acertadamente explicam Peixoto e Silva (2019, p. 89), o que ocorre, na prática, é a execução de tarefas comandadas pelos desenvolvedores, sendo então possível aos computadores, a partir de um padrão neles inserido ou reconhecido, lidar com problemas que exigem conhecimentos do mundo real, o que possibilita a tomada de decisões que aparentam humanas.

Esse padrão denomina-se *algoritmo*, o qual é a fórmula lógica constante dos programas de IA que determina o processamento de dados a ser efetuado ou, conforme expõem Peixoto e Silva (2019, p. 71), com base em Hanke (2018), algoritmo é o conjunto de regras que definem uma sequência de operações.

Assim, conceitua-se IA como *software* ou *programa* dotado de *algoritmo* com elevada capacidade de *processamento de dados*.

Desse conceito e da constatação de que considerável parte das asserções feitas em relação à IA, sua capacidade e implicações, são frequentemente muito mais impressões aparentes do que verdades essenciais, decorre que os procedimentos praticados pela máquina são passíveis de *conhecimento e verificação humana*. Em outras palavras, é sempre possível – embora possa ser extremamente trabalhoso e dispendioso – o domínio sobre os parâmetros utilizados pela IA. Dessa forma, na relação entre o ser humano e a máquina, o *controle* cabe ao *ser humano* e não vice-versa.

Em outras palavras, conforme explicam Nunes e Marques (2018, p. 424), os desenvolvedores selecionam as informações ou dados que alimentarão o sistema para torná-lo apto a fornecer soluções para determinadas situações-problema. A elaboração dos algoritmos e a escolha dos dados neles inseridos revela a atuação das máquinas dentro do especificado por seres humanos.

Não obstante, um algoritmo pode ser construído de tal modo que contenha instruções que determinam o modo pelo qual o próprio algoritmo, ao realizar as mesmas tarefas de combinação e parametrização de dados, retorne *outputs* mais adequados (respostas mais qualificadas para o problema apresentado) a cada novo processamento do programa (VALENTINI, 2018, p. 59). Esse processo é o que se chama de aprendizado de máquina ou *machine learning*, em que se tem a aparente impressão de que a máquina “aprende” com a

experiência. Todavia, por mais complexo que se torne o processamento de dados em virtude do elevado número de códigos e parâmetros, a atuação das máquinas se dá dentro da programação humana, porquanto é a intervenção humana que determina a forma como os dados inseridos (*inputs*) serão processados e externados (*outputs*).

Isto é, os comportamentos se tornam relativamente autônomos, mas, em verdade, apenas dentro da constituição algorítmica e de sua base de dados. No caso do *machine learning*, embora a máquina seja capaz de aperfeiçoar seu banco de dados, a própria busca por novas informações ocorrerá dentro dos parâmetros estabelecidos por quem a programou (NUNES; MARQUES, 2018, p. 05).

Cabe destacar que os dados primários oferecidos ao equipamento e utilizados em seu treinamento servem como base para as suas conexões, as quais são sucessivamente processadas e combinadas de modo a alcançar níveis de generalização em que se viabiliza a abstração dos dados específicos inicialmente usados. Markou e Deakin (2019, p. 17-18) explicam que quaisquer dados inseridos nas máquinas produzirão, invariavelmente, resultados parciais, limitados às possibilidades de análise e ao universo de ação de cada código ou algoritmo. O grau de generalização e abstração torna-se proporcionalmente mais elevado à medida em que o alto número de parâmetros forma uma *rede* de “aprendizado profundo” (*deep learning*), que permite acesso a uma quantidade abrangente de dados fundamentais.

Entretanto, relevante ponto a ser lembrado é a existência de margem de erro da IA no processamento de dados, na sua organização segundo padrões, assim como a possibilidade de nela haver *viés algorítmico* (*algorithm bias*), ou seja, enviesamento nas respostas resultantes do processamento de dados pelos algoritmos de IA.

Contudo, deve-se notar que, sendo os algoritmos humanamente programados e os dados neles inseridos humanamente selecionados, o viés não é algorítmico, mas, sim, humano. O enviesamento resulta da subjetividade no desenvolvimento do *software*, na escolha dos dados a serem utilizados, pois essa seleção reflete objetivos, prioridades, concepções, valores pessoais dos desenvolvedores da IA (NUNES, MARQUES, 2018, p. 424). Tais concepções e valores pessoais podem representar preconceitos, expressar posições intolerantes e discriminatórias, que então distorcem as respostas fornecidas pela máquina. Entretanto, percebe-se que, ao se fundar na subjetividade do desenvolvedor, o viés algorítmico é vício *humano*, do indivíduo que incute no programa de IA valores e crenças naturalizados da sociedade em que se formou ou ainda, como explica Toledo (2021, p. 82), representa

a codificação de concepções discriminatórias humanas, uma vez que os dados

empregados pelos programas de IA, os pesos a eles atribuídos, as combinações entre eles feitas, as relações de implicação e inferência entre eles estabelecidas são todos padrões decisórios utilizados para a tomada de decisões humanas.

Finalmente, deve-se destacar que vieses algorítmicos não apenas *reproduzem* preconceitos e prevenções, mas os *potencializam*, uma vez que os resultados discriminatórios por eles gerados são majorados na mesma proporção que a capacidade combinatória de padrões e processamento de dados da IA é superior à capacidade humana. Ressalte-se que se vieses algorítmicos presentes em programas de IA utilizados no âmbito privado das relações particulares são condenáveis e devem ser corrigidos, a gravidade dos prejuízos por eles trazidos e a urgência da necessidade de sua correção são notoriamente maiores no caso de sua utilização pelo Estado. Em verdade, “a utilização de programas de IA com viés algorítmico pelo Poder Público significa a *institucionalização* do desvio, a oficialização do tratamento discriminatório, enfim, a legalização do ilegal” (TOLEDO, 2021, p. 82-83).

Neste momento, clara fica a relevância e indispensabilidade do Direito no tratamento da IA, uma vez que a correção desses vieses faz-se pela subordinação da produção da IA às diretrizes éticas positivadas nas normas jurídicas, ou seja, a correção de vieses algorítmicos se faz pela regulamentação legal da matéria, segundo os parâmetros regentes da vida social. Os parâmetros valorativos máximos em dimensão nacional estão nas normas constitucionais, sendo marcadamente dispostos nos direitos fundamentais vigentes no Estado Democrático de Direito. Em âmbito internacional, esses parâmetros estão expressos nos direitos humanos.

## CONCLUSÕES

Esta pesquisa dedicou-se inicialmente a elaborar conceito cientificamente válido para o que se denomina de *inteligência artificial*, IA, para, a partir dele, aprofundar no estudo de suas capacidades e então compreender a sua possibilidade de aplicação dentro do Direito.

Conceituou-se IA como *programa computacional* dotado de *algoritmo* com elevada capacidade de *processamento de dados*. Uma das primeiras conclusões que se extraem de tal conceito é a de que a IA é incapaz de pensar por si só, mas rigorosamente apenas trabalha com códigos e dados inseridos por seu desenvolvedor. Desse fato, seguem-se então conclusões adicionais:

- a) Os algoritmos elaborados e os dados selecionados para serem neles inseridos são permeados de carga valorativa dos seus desenvolvedores, porquanto são produção humana;

- b) Consequentemente, vieses cognitivos humanos são reproduzidos nos algoritmos e dados com os quais trabalha a IA, resultando nos chamados *vieses algorítmicos*;
- c) Os vieses algorítmicos não apenas reproduzem os vieses cognitivos humanos, mas os *potencializam* exponencialmente, devido à elevada capacidade e rapidez da máquina no processamento de dados;
- d) Entretanto, da mesma forma que, apesar da complexidade e dificuldade envolvidas, vieses cognitivos humanos são passíveis de identificação e correção pelo próprio ser humano, também os vieses algorítmicos, apesar da complexidade e sofisticação dos códigos desenvolvidos, são passíveis de verificação e correção humanas;
- e) A correção de vieses algorítmicos vincula-se às normas jurídicas vigentes, dentre as quais sobressaem os direitos fundamentais e as normas estruturantes do Estado Democrático de Direito.

## REFERÊNCIAS

MARKOU, Christopher; DEAKIN, Simon. Ex Machina Lex: The Limits of Legal Computability. 2019. Disponível em: <[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3407856](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3407856)>. Acesso em: 14 nov. 2020.

NUNES, Dierle; MARQUES, Ana Luiza Pinto Coelho. A inteligência artificial e o direito processual: Vieses algorítmicos e os riscos de atribuição de função decisória às máquinas. Revista dos Tribunais Online, v. 285, p. 421-447, 2018. Disponível em: [https://www.academia.edu/37764508/INTELIG%C3%8ANCIA\\_ARTIFICIAL\\_E\\_DIREITO\\_PROCESSUAL\\_VIESES\\_ALGOR%C3%8DTMICOS\\_E\\_OS\\_RISCOS\\_DE\\_ATRIBUI%C3%87%C3%83O\\_DE\\_FUN%C3%87%C3%83O\\_DECIS%C3%93RIA\\_%C3%80S\\_M%C3%81QUINAS\\_-\\_Artificial\\_intelligence\\_and\\_procedural\\_law\\_algorithmic\\_bias\\_and\\_the\\_risks\\_of\\_assignment\\_of\\_decision-making\\_function\\_to\\_machines](https://www.academia.edu/37764508/INTELIG%C3%8ANCIA_ARTIFICIAL_E_DIREITO_PROCESSUAL_VIESES_ALGOR%C3%8DTMICOS_E_OS_RISCOS_DE_ATRIBUI%C3%87%C3%83O_DE_FUN%C3%87%C3%83O_DECIS%C3%93RIA_%C3%80S_M%C3%81QUINAS_-_Artificial_intelligence_and_procedural_law_algorithmic_bias_and_the_risks_of_assignment_of_decision-making_function_to_machines). Acesso em: 31 out. 2021.

PEIXOTO, Fabiano Hartmann; SILVA, Roberta Zumblick Martins da. Inteligência Artificial e Direito. Curitiba: Alteridades, 2019. (Coleção Direito, Racionalidade e Inteligência Artificial).

TOLEDO, Cláudia; ALVES, Michelle. Inteligência Artificial Aplicada ao Direito. I Seminário de Pesquisa do Observatório de Práticas Jurídicas da UFERSA, mai. 2021.

TOLEDO, Cláudia. Inteligência Artificial e sua aplicabilidade em decisões judiciais. In: PEIXOTO, Fabiano Hartmann (Org.). Inteligência Artificial: Estudos de Inteligência Artificial. Curitiba: Alteridade, 2021, p. 57-90.

VALENTINI, Rômulo Soares. Julgamento por computadores? As novas possibilidades da juscibernética no século XXI e suas implicações para o futuro do Direito e do Trabalho dos Juristas. Tese (Doutorado em Direito). Faculdade de Direito, Universidade Federal de Minas

Gerais, 2018.