

INTELIGÊNCIA ARTIFICIAL JUDICIAL E EXPLICABILIDADE: AS RAZÕES PÚBLICAS E A JUSTIÇA “EX MACHINA”

JUDICIAL ARTIFICIAL INTELLIGENCE AND EXPLAINABILITY: PUBLIC REASONS AND “EX MACHINA” JUSTICE

Sérgio Rodrigo de Pádua

Resumo

O presente artigo tem por objetivo analisar aspectos iniciais do direito à explicabilidade das decisões judiciais tomadas com auxílio de inteligência artificial. A solução proposta é o desenho, em relação aos aspectos básicos, de um direito fundamental à explicabilidade pautado no pluralismo democrático decorrente das razões públicas. A abordagem da pesquisa foi multidisciplinar, eis que foram objeto de análise os diplomas normativos que vêm sendo produzidos sobre o tema, categorias teóricas da justiça política de John Rawls, e de estudos científicos sobre os limites e as possibilidades da explicabilidade das decisões tomadas por inteligência artificial. Este estudo se pauta em pesquisa qualitativa, mediante análise exploratório-descritiva inerente ao Direito e ao Ciência da Computação. Os resultados demonstram: a necessária interseção entre a teoria da justiça mediante as razões públicas e as técnicas de inteligência artificial explicável (xAI), a fim de se garantir a legitimidade democrática dos sistemas de IA judicial; e a definição de direito fundamental à explicabilidade estrita da decisão judicial tomada com auxílio de IA. Portanto, a IA de auxílio à decisão judicial (pensamento híbrido humano-máquina) já funciona com as limitações que a tecnologia impõe, cabendo desde já reconhecer a explicabilidade da IA jurídica como um direito necessário para uma sociedade plural e justa.

Palavras-chave: Razões públicas; Teoria da justiça; Inteligência artificial explicável; xAI; Explicabilidade.

Abstract

The purpose of this article is to analyze the initial aspects of the right to explainability of judicial decisions by artificial intelligence. The proposed solution is the design of a fundamental right of explainability. The right to explainability is based on democratic pluralism arising from public reasons. The research approach was multidisciplinary.

The research analyzed the normative acts produced on the subject, theoretical categories of political justice in Rawls' think), and scientific studies on the limits and possibilities of explainability in artificial intelligence. This study based on qualitative research, through exploratory-descriptive analysis about Law and Computer Science. The results demonstrate: the necessary intersection between the theory of justice (approach to public reasons) and the techniques of explainable artificial intelligence (xAI) to guarantee the democratic legitimacy of judicial AI systems; and the definition of a fundamental right to explainability of the judicial decision by AI. Therefore, Artificial Intelligence to support the court decision (human-machine hybrid thinking) works with limitations imposed by technology. Thus, the explicability of legal AI needs to be recognized as a right. This will enable plurality and justice in society.

Keywords: Public reasons; Theory of justice; Explainable artificial intelligence; xAI; Explainability.

INTRODUÇÃO

Falar de inteligência artificial deixou de ser tema de *nerd* e de filmes de ficção científica, pois já se pode vislumbrar as primeiras facetas daquilo que tem se convencionado chamar de revolução 4.0¹, um contexto fático de relevância histórica, econômica e social que cada vez mais hibridiza máquinas e humanidade num só corpo. A era da Inteligência Artificial só está começando. Isso traz a reflexão sobre quais características hoje adjetivadas como “humanas” restarão num mundo de uma sociedade mergulhada em tecnologia.

Os céticos e os enamorados pela IA estão em toda parte, e não seria o Direito o único a passar imune à essa *AI storm*. Nesse sentido, o ser humano busca inocular justiça nos algoritmos computacionais que animam as máquinas, muito disso porque a humanidade ainda tem seus sonhos de brincar de deus através da IA.

Falar de inteligência artificial e seus impactos jurídicos e filosóficos é um caminhar pela sala dos desejos da humanidade. Os sonhos de criação de inteligências

¹ Schwab, Klaus. A Quarta Revolução Industrial. São Paulo: Edipro, 2018, edição do Kindle, posição 2425.

artificiais amistosas à humanidade e os pesadelos sobre uma IA maléfica habitam o imaginário cotidiano. Conviver com algo próximo da figura amistosa de *C-3PO* de *Star Wars* parece mais próximo do que conhecer a maquinaria desumanizante de tantas outras obras de ficção.² Para isso o Direito deve se debruçar sobre as diversas questões que os avanços da inteligência artificial lhe causam.

Dessa maneira, a hipótese estudada no presente artigo se refere à possibilidade de entrelaçamento entre os conceitos de razão pública, tipicamente advinda do consenso sobreposto da teoria da justiça política, e as técnicas de inteligência artificial explicável (xAI) para a definição do direito à explicabilidade às decisões judiciais auxiliadas por IA.

A presente pesquisa é de cunho qualitativo e mediante método exploratório-descritivo, por meio de análise interdisciplinar entre Teoria da Justiça, Direito Constitucional e da Ciência da Computação, de maneira a delinear os aspectos relevantes da relação entre a concepção pluralista de justiça de John Rawls e a construção do direito à explicabilidade das decisões tomadas com auxílio de sistemas de inteligência artificial jurídica.

O primeiro capítulo faz um panorama da construção da ideia de explicabilidade das decisões auxiliadas por inteligência artificial. O segundo capítulo traz definições da interação entre as razões públicas e a Ciência da Computação para a caracterização do direito à explicabilidade. Já o terceiro capítulo emoldura a concepção de explicabilidade, inerente às técnicas de inteligência artificial explicável (xAI), à ideia de justiça através da construção das decisões judiciais auxiliadas por IA que levem em consideração o consenso contraposto advindo das razões públicas.

A solução proposta é o desenho inicial, em relação aos aspectos básicos, de um direito fundamental à explicabilidade pautado no pluralismo democrático decorrente das razões públicas.

Destaque-se que o artigo é um recorte (necessário pelo escopo presente) de uma pesquisa maior da relação entre decisão judicial e inteligência artificial.

Indicadas as bases do estudo, passa-se aos tópicos abordados.

² Bostrom, Nick. *Superinteligência: Caminhos, perigos, estratégias*. Rio de Janeiro: Darkside Books, 2014, edição do Kindle, p. 8864-8877.

1 INTELIGÊNCIA ARTIFICIAL JUDICIAL: A CONSTRUÇÃO DA IDEIA DE EXPLICABILIDADE

O conhecimento não é feito só com o “velho”, é fruto de uma obra de desconstrução para uma reconstrução. O pensamento de Jürgen Habermas³, por exemplo, mostra isso. A ciência vive de uma arquitetura construída “*sobre os ombros de gigantes*” (Isaac Newton)⁴, mas às vezes, de tempos em tempos, passa por grandes transformações.

As obras de ficção são uma fagulha inicial ao debate, um estímulo ao pensamento acerca de quais seriam os pressupostos jurídico-filosóficos para a constituição de uma civilização humano-máquina, na busca de explicar a simbiose entre o pensamento humano e suas características típicas (pensamento rápido através de heurísticas e pensamento reflexivo)⁵ em interação com um aparato de soluções decisórias trazidas pelas técnicas de inteligência artificial.⁶ Assim, as preocupações com aspectos da justiça numa interação humano-máquina, ainda que tenham se iniciado em obras de ficção científica, cada vez mais permeiam a realidade do mundo atual.

As três principais posições perante o problema têm sido o alarmismo apocalíptico (exemplificado no pensamento de Yuval Harari⁷), o desalento com o novo (que tem tomado a psique de grande parte dos juristas⁸) e o arrebatamento tecnológico (de parte da comunidade científica⁹). Todavia, é necessário um pouco de realidade como temperamento às posições referidas.

³ Habermas, Jürgen. Teoria do Agir Comunicativo: Racionalidade da Ação e Racionalidade Social. Vol. 1. São Paulo: WMF Martins Fontes, 2012, p. 205-206.

⁴ “*If I have seen further it is by standing on the shoulders of Giants*” (NEWTON, Isaac. “Letter from Sir Isaac Newton to Robert Hooke”, 1675. Disponível em: <https://digitallibrary.hsp.org/index.php/Detail/objects/9792>. Acesso em: 06/04/2021).

⁵ Kahneman, Daniel. Rápido e Devagar: Duas formas de pensar. Rio de Janeiro: Objetiva, 2012, edição do Kindle, posição 613

⁶ Taulli, Tom. Introdução à Inteligência Artificial: Uma abordagem não técnica. São Paulo: Novatec, 2020, p. 187-190.

⁷ Harari, Yuval Noah. Homo Deus: Uma breve história do amanhã. Tradução: Paulo Geiger. São Paulo: Companhia das Letras, 2016, p. 313.

⁸ Susskind, Richard. Online Courts and the Future of Justice. Oxford: Oxford University Press, edição do Kindle, p. 44.

⁹ Howard, Philip N. Pax Technica: How the Internet of Things May Set Us Free or Lock Us Up. New Haven: Yale University Press, 2015, edição do Kindle, p. 233-234.

O desvelar do novo insiste em ficar no futuro¹⁰, longe dos olhos curiosos do presente. Todavia, já existe muito presente para ser trabalhado. Nessa linha, estudo realizado pela Fundação Getúlio Vargas, através de seu Centro de Inovação, Administração e Pesquisa do Judiciário, indicou que existem sessenta e quatro projetos¹¹ de inteligência artificial judicial em quarenta e sete Tribunais do Brasil¹², muitos destes projetos já fase de implementação. Vários dos projetos de IA judicial, como, por exemplo, os sistemas *Victor* (STF), *Sócrates* (STJ), *Sinapse* (TJRO), *Sigma* (TRF3), *Sugestão de Modelos de Minutas* (TRF4) e *JUDI* (TJSP) impactam diretamente no processo de decisão dos julgadores, o que denota a relevância do debate acerca do presente tema.

Levando-se em conta esta realidade, o atual sentimento de estranhamento com a tecnologia se constrói a partir da crise com aquilo que ainda não existe (mas poderá existir), mesclado com o espanto com aquilo que já é. Essa a fotografia do momento da relação entre decisão judicial e IA. Todavia, olhares criteriosos são necessários, pois a decisão judicial no entrelaçamento juiz-IA é fenômeno que traz riscos e potencialidades antes não experimentados, no plano dos riscos emergentes se destacam os vieses algoritmos¹³, os quais têm atraído bastante atenção da comunidade jurídica (embora os vieses já habitassem a mente dos julgadores muito antes dos computadores existirem¹⁴), e a maior opacidade das decisões auxiliadas por IA, sendo este último o ponto de tensão no qual o presente estudo se foca.

¹⁰ Kurzweil, Ray. *A Singularidade Está Próxima: quando os humanos transcendem a biologia*. Tradução: Ana Goldberger. São Paulo: Iluminuras, 2018, p. 27.

¹¹ Os principais são: *Victor* no STF; *Athos*, *Sócrates*, *E-Juris* e *TUA* no STJ; *BEM-TE-VI* no TST; Plataforma *Sinapses* no CNJ; *Robô Secor*, *Banco de Sentenças*, *SIB*, *ALEI* e *Projeto Execução Célere* no TRF1; *Atendente Virtual* no TRF2; *SINARA*, *SIGMA* e *Prevenção* no TRF3; *Classificação de Temas na Vice-Presidência* e *Turmas Recursais, Análise de Assunto dos Processos, Triagem Automática de Processos a Partir da Petição Inicial* e *Sugestão de Modelos de Minutas* no TRF4, *JULIA* no TRF5; *LEIA* em seis Tribunais de Justiça diversos (TJAC, TJAL, TJAM, TJCE, TJMS e TJSP); *Hércules* no TJAL; *Queixa Cidadã* no TJBA; *Hórus* e *Ámon* no TJDF; *IA332* no TJGO; *ELIS* no TJPE; *Sinapse* no TJRO; *SCRIBA* e *Mandamus* no TJRR; *JUDI* no TJSP; *MINERJUS* no TJTO; *Inteligência Artificial e Eficiência do Judiciário* no TRT1; *Clusterização de Processos* no TRT4; *GEMINI* em quatro Tribunais Regionais do Trabalho diversos (TRT5, TRT7, TRT15 e TRT20); *B.I. TRT 11* no TRT11; e *CONCILIA JT* no TRT12.

¹² Salomão, Luis Felipe (coord.) et al. *Inteligência Artificial: Tecnologia Aplicada à Gestão dos Conflitos no Âmbito do Poder Judiciário Brasileiro*. FGV Conhecimento: Centro de Inovação, Administração e Pesquisa do Judiciário, 2020, p. 26.

¹³ Yapo, Adrienne; Weiss, Joseph. "Ethical Implications of Bias in Machine Learning", em *Proceedings of 51st Hawaii International Conference on System Sciences*, Waikoloa Village, 2018, p. 5365-5372.

¹⁴ Sunstein, Cass R. "Algorithms, Correcting Biases", em *Social Research: Algorithms*, v. 86, n. 2, 2019, New York.

Todavia, para ter condições de escolher o ser humano deve efetivamente conhecer as decisões que sobre ele são tomadas na sociedade democrática da qual ele faz parte¹⁵, o que inclui as decisões auxiliadas por sistemas de IA judicial. Dessa forma, o desenvolvimento de inteligência artificial explicável (xAI) é necessário para que os padrões democráticos pautados na transparência das decisões públicas sejam respeitados.

Nessa linha, a inteligência artificial explicável (xAI) envolve esforços para explicar como um modelo de IA chegou ao resultado (*output*) verificado¹⁶ ou ainda para auxiliar os humanos na interpretação deste.¹⁷

Nesse sentido, em relação às decisões judiciais auxiliadas por inteligência artificial, a teoria das razões públicas como formadoras das decisões de repercussão na sociedade foi trazida por John Rawls em sua segunda fase de pensamento acerca da teoria da Justiça, fase na qual o autor se pautou na construção de uma ideia de justiça política.¹⁸ A concepção de justiça política traz consigo uma abertura para diversas possibilidades de justiça que podem ser tidas como “razoáveis”¹⁹ na construção de um sistema constitucional democrático marcado pelo pluralismo como característica.²⁰

É justamente tendo por base o natural pluralismo que, numa sociedade democrática, a explicabilidade das decisões tomadas por autoridades estatais ganha a importância de ser o meio apto a dar a possibilidade de publicidade, de transparência e, mais do que isso, de *accountability*²¹ sobre referidas decisões.

¹⁵ Rawls, John. *Justiça e Democracia*. São Paulo: Martins Fontes, 2000, p. 216.

¹⁶ “xAI encompasses a range of efforts to explain—or help humans interpret—how a particular machine learning model reached its conclusion. The concept of an explanation here ‘has come to refer to providing insight into the internal state of an algorithm, or to human-understandable approximations of the algorithm.’ xAI provides a variety of benefits: It can foster trust between humans and the system, identify cases in which the system appears to be biased or unfair, and bolster our own knowledge of how the world works” (Deeks, Ashley. “The Judicial Demand for Explainable Artificial Intelligence”, em *Columbia Law Review*, v. 119, n. 7, New York, p. 1834).

¹⁷ Gilpin, Leilani H. et. al. “Explaining Explanations: An Overview of Interpretability of Machine Learning”, em *2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)*, 2018, Turin, p. 80-89.

¹⁸ Gargarella, Roberto. *As Teorias da Justiça Depois de Rawls: Um breve manual de filosofia política*. São Paulo: WMF Martins Fontes, 2008, p. 227-228.

¹⁹ Rawls, John. *Justiça e Democracia*, p. 254-255.

²⁰ *Ibid.* Prefácio, p. X e XI.

²¹ Wagner, Ben. “Accountability by design in technology research”, em *Computer Law & Security Review*, v. 37, 105398, ISSN 0267-3649, 2020, Oxford.

Especificamente em relação ao Poder Judiciário, com a recente inclusão dos sistemas de IA judicial como novos atores no auxílio à tomada de decisão, a necessidade de explicabilidade toma corpo. A explicabilidade é trazida como um elemento da transparência das decisões tomadas por IA através de diversas iniciativas, como na Recomendação 449 da OCDE²², na *First Version of a Draft Text of a Recommendation on the Ethics of Artificial Intelligence*²³ da UNESCO, nos *Ethical Guidelines for Trustworthy AI*²⁴ da União Europeia, no art. 20 da Lei Geral de Proteção de Dados (Lei 13.709/2018) e na Resolução 332/2020 do Conselho Nacional de Justiça.²⁵

O art. 19 da Resolução 332/2020 do CNJ prevê que os sistemas de inteligência artificial utilizados em auxílio à elaboração de decisões judiciais “observarão, como critério preponderante para definir a técnica utilizada, a explicação dos passos que conduziram ao resultado”.²⁶ A transparência no uso da IA judicial, segundo o art. 8º, VI, da Resolução 332/2020 do CNJ, tem como um de seus elementos “o fornecimento de explicação satisfatória e passível de auditoria por autoridade humana quanto a qualquer proposta de decisão apresentada pelo modelo de Inteligência Artificial”.²⁷ A norma do CNJ desenvolveu sua ideia de explicabilidade aparentemente com inspiração no **princípio da transparência judicial** previsto nos *Asilomar AI Principles*, segundo o qual “qualquer envolvimento de um sistema autônomo na tomada de decisões judiciais deve fornecer uma explicação satisfatória, passível de auditoria por uma autoridade humana competente”.²⁸

As novas abordagens referidas vêm ao encontro da exigência de explicabilidade das decisões judiciais que tradicionalmente permeia a exposição da motivação decisória, resumidamente composta pelos motivos de fato e pelos motivos

²² OCDE. “Recommendation 0449, of 21 May 2019: Recommendation of the Council on Artificial Intelligence”. Disponível em: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>. Acesso em: 04/08/2020.

²³ UNESCO. “First version of a draft text of a recommendation on the Ethics of Artificial Intelligence”. 15 May 2020. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000373434>. Acesso em: 25/07/2020.

²⁴ Comissão Europeia. “Ethics Guidelines for Trustworthy AI”. 08 April 2019. Disponível em: https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419. Acesso em: 25/07/2020.

²⁵ Brasil. Conselho Nacional de Justiça. Resolução 332, de 21 de agosto de 2020. Dispõe sobre a ética, a transparência e a governança na produção e no uso de Inteligência Artificial no Poder Judiciário e dá outras providências.

²⁶ Id.

²⁷ Id.

²⁸ Future of Life Institute. *Asilomar AI Principles*. 2017. Disponível em: <https://futureoflife.org/ai-principles>. Acesso em: 05/08/2020.

de direito (normas, posicionamentos doutrinários, precedentes judiciais e sua interpretação), o que passa pela análise das premissas utilizadas na decisão, na forma exigida no art. 489 do Código de Processo Civil e no art. 93, IX, da Constituição Federal, sendo que tais premissas advêm de uma análise mais aprofundada dos pressupostos consensuais²⁹ que constituem a decisão. Nessa linha, mediante processo de atribuição³⁰ pode-se inferir um direito fundamental à explicabilidade das decisões judiciais auxiliadas por inteligência artificial.

Contudo, a explicabilidade dos modelos de inteligência artificial é cercada de dificuldades técnicas que surgem desde os modelos algorítmicos utilizados na programação, como as abordagens procedimental³¹, de representação do conhecimento³² e de aprendizado de máquina (*machine learning*)³³ que podem implementadas.

Muito daquilo que na atualidade se entende como simples programação (programação orientada a objetos, por exemplo) já foi tida, num passado recente, como “inteligência artificial”. IA não “é mágica ou ficção científica, mas ciência, engenharia e matemática”³⁴. É curioso como os juristas querem definir tudo, inclusive o que é (ou não é) inteligência artificial, sendo que o conceito de inteligência artificial apenas depende da existência de grau de autonomia decisória inerente aos algoritmos do programa de computador³⁵, grau de autonomia que pode ser maior ou menor.

Nos sistemas de representação do conhecimento destaca-se na área de programação jurídica o desenvolvimento de sistemas especialistas³⁶, bem como tem cada vez mais ganhado espaço a aplicação de sistemas baseados em aprendizado de máquina (*machine learning*).³⁷

²⁹ Habermas, Jürgen. Op. cit., p. 45.

³⁰ Alexy, Robert. Teoria dos Direitos Fundamentais. 2ª ed. São Paulo, 2011, p. 74.

³¹ Rover, Aires José. Informática no Direito: Inteligência Artificial. Curitiba: Juruá, 2001, p. 66-67.

³² Ashley, Kevin. Artificial Intelligence and Legal Analytics: New tools for law practice in digital age. Cambridge: Cambridge University Press, 2017, edição do Kindle, p. 81.

³³ Lage, Fernanda de Carvalho. Manual de Inteligência Artificial no Direito Brasileiro. Salvador: Editora Juspodivm, 2021, p. 70-71.

³⁴ Russell, Stuart; Norvig, Peter. Inteligência Artificial. 3ª ed. Rio de Janeiro: Elsevier, 2013, p. 27.

³⁵ Dewey, Daniel; Russel, Stuart; Tegmark, Max. “Research Priorities for Robust and Beneficial Artificial Intelligence”, em AI Magazine, v. 36, n. 4, 2015.

³⁶ Rover, Aires José. Op. cit., p. 217-220.

³⁷ Rover, Aires José. “Machine Learning no Poder Judiciário: uma biblioteca temática, ano 2020”, em Hartmann Peixoto, Fabiano (Org.). Inteligência Artificial: estudos de inteligência artificial. Vol. 4. Curitiba: Alteridade, 2021, p. 20-21.

Especificamente em relação às redes neurais de aprendizado profundo, modelos que simulam uma rede extremamente conectada de neurônios em diversas camadas (*layers*), estas têm vantagens e desvantagens.³⁸ Maior exemplo do desenvolvimento de modelos *deep learning* é a *GPT-3 (Generative Pre-training Transformer 3)*³⁹ da *OpenAI*⁴⁰, a qual é uma IA de processamento de linguagem natural (*NLP*), com pretensão de aplicabilidade geral (*task agnostic*), baseada em pesquisas em redes neurais profundas⁴¹. Os dados a respeito do desenvolvimento do *GPT-3* foram publicados pela desenvolvedora do sistema⁴², o que possibilita a pesquisa.

A *API* do *GPT-3* está em fase de testes, todavia já tem sido aplicada em problemas jurídicos conjuntamente com outros modelos de IA, como no caso da *CARA A.I.* da empresa *Casetext*⁴³, por exemplo, sistema hibridizado (*GPT-3* e IA com desenvolvimento para tarefas específicas) que é utilizado para pesquisa de jurisprudência, leis e doutrina, bem como é capaz de gerar texto com o emprego de conceitos jurídicos, o que tem auxiliado os profissionais do Direito⁴⁴.

Contudo, a utilização de sistemas baseados em tecnologia *deep learning*, embora seja de grande valia no auxílio a advogados, por exemplo, traz consigo o risco de opacidade⁴⁵ caso seja utilizada como sistema auxiliar para a tomada de decisões

³⁸ MIT Technology Review. “Por que a GPT-3 é o melhor e o pior da IA atualmente”. 01/04/2021. Disponível em: <https://mittechreview.com.br/por-que-a-gpt-3-e-o-melhor-e-o-pior-da-ia-atualmente>. Acesso em 08/04/2021.

³⁹ Brown, Tom B. et al. “Language Models are Few-Shot Learners”. OpenAI, San Francisco, arXiv:2005.14165. Disponível em: <https://arxiv.org/pdf/2005.14165.pdf>. Acesso em 05/04/2021.

⁴⁰ OpenAI. OpenAI API. Disponível em: <https://openai.com/blog/openai-api/>. Acesso em 08/04/2021.

⁴¹ Desde o lançamento do *GPT-3*, em 2020, houve grande impacto na comunidade científica. O *GPT-3* é uma IA de processamento de linguagem natural (*NLP*). O *GPT-3* é uma IA *task agnostic* via rede neural profunda (*DNN*) com 175 bilhões de parâmetros. O treinamento inicial do sistema *deep learning* usou o seguinte *dataset* em número de *tokens*: *Common Crawl* (410 bilhões – peso 60%), *WebText2* (19 bilhões), *Books1* (12 bilhões), *Books2* (55 bilhões) e *Wikipedia* (3 bilhões). A maior base de dados é o *Common Crawl*, que é um *crawler* de leitura de conteúdos existentes em toda a internet. Em relação à língua dos textos utilizados no treinamento tem a seguinte distribuição por número de palavras: a) inglês – 92,64708% – 181.014.683.608; b) francês – 1,81853% – 3.553.061.536; c) alemão – 1,46937% – 2.870.869.396; d) espanhol – 0,77289% – 1.510.070.974; e) italiano – 0,60793% – 1.187.784.217; f) português – 0,52483% – 1.025.413.869; e outros 112 idiomas em menores percentuais.

⁴² Brown, Tom B. et al. Op. cit.

⁴³ Casetext. *CARA A.I.* Disponível em: <https://casetext.com/cara-ai>. Acessado em 28/03/2021.

⁴⁴ Não há ainda uma visão de substituição dos operadores do Direito pela IA referida, pois esta tem menor acurácia que os humanos na atribuição de significado, uma vez que 76% no *GPT-3 Small* e 52% no *GPT-3*. A equipe de pesquisa da *OpenAI*, mediante os dados de desempenho fornecidos, admite que o modelo completo (*GPT-3*) com 175 bilhões de parâmetros tem baixa acurácia na tarefa de interpretar o real significado das frases geradas (BROWN, Tom B. et al. Op. cit.).

⁴⁵ Deeks, Ashley. Op. cit., p. 1829.

judiciais, uma vez que existem grandes dificuldades técnicas em relação à possibilidade de explicabilidade das saídas (*outputs*) desta classe de modelos de IA⁴⁶, o que impacta na transparência inerente aos padrões mínimos de justiça que se buscam desenvolver numa sociedade democrática.

2 INTELIGÊNCIA ARTIFICIAL JUDICIAL EXPLICÁVEL: DIREITO À EXPLICABILIDADE E RAZÕES PÚBLICAS

A explicabilidade dos resultados do processamento advindo de algoritmos de IA passa necessariamente pela ideia de *accountability*, sendo a explicabilidade um aprimoramento desta abordagem.

Dessa maneira, a *accountability* no desenvolvimento de modelos de IA judicial deve passar pela transparência no processo construção do modelo algorítmico⁴⁷, sendo que a explicabilidade advinda da transparência no procedimento de funcionamento da IA é abordagem mais efetiva que se apresenta⁴⁸, a qual pode ser dividida em fases de garantia da explicabilidade de modelos de IA⁴⁹, o que consiste na definição do objetivo (primeira fase)⁵⁰, na escrita do código algorítmico (*coding* – segunda fase)⁵¹ e na implementação do modelo (terceira fase).⁵² Dessa forma, pode-se compreender que a explicabilidade do processo de construção de modelos de IA tem pressupostos característicos que são distintos da explicabilidade das saídas (*outputs*) e resultados (*outcomes*) propiciados por um modelo de IA.

Nessa linha, no processo de construção das possibilidades de explicabilidade dos modelos de inteligência artificial judicial deve existir deliberação consciente acerca dos objetivos e das prioridades adotadas pelos desenvolvedores, parâmetros que devem ser avaliáveis na esfera pública.⁵³ Para além disso, no momento em que a IA

⁴⁶ Pasquale, Frank. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge: Harvard University Press, 2015, edição do Kindle, p. 216-217.

⁴⁷ Licht, K. de Fine; Licht, J. de Fine. *Artificial intelligence, transparency, and public decision-making*. *AI & Society*, Mar. 2020.

⁴⁸ Id.

⁴⁹ Id.

⁵⁰ Id.

⁵¹ Id.

⁵² Id.

⁵³ Id.

é implementada o código-fonte e os registros de como a IA é utilizada necessitam ser disponibilizados publicamente, conforme foi regulado pelo Conselho Nacional de Justiça mediante o depósito dos modelos de IA judicial no Sinapses (art. 10, III, do art. 332/2020 do CNJ), o qual é solução computacional para “armazenar, testar, treinar, distribuir e auditar modelos de Inteligência Artificial” (art. 3º, III, do art. 332/2020 do CNJ).

Dessa maneira, a tomada de decisão judicial com o auxílio de IA pode ser mais célere e eficiente sem que necessariamente haja perda da qualidade deliberativa, uma vez que a habilidade de chegar a uma decisão e implementá-la nos algoritmos é importante na tomada de decisões através de IA, desde que garantido o posterior controle e a supervisão por humanos (jugador e sua equipe técnica)⁵⁴, eventuais vieses e ou saídas incorretas (sem acurácia)⁵⁵ podem ser tratados antes de ser efetivamente adotada a decisão sugerida pela IA judicial, devendo o erro da IA ser registrado no sistema (art. 25, Parágrafo único, V, da Res. 332/2020) para que exista a transparência da decisão tomada.

Neste ponto, a não realização do registro serviria como vulnerabilidade para a burla à proposta decisória advinda da IA sem a necessidade de demonstração dos motivos de tal superação, indicando-se em quais pontos e por quais motivos a decisão puramente humana do julgador (agora auxiliado apenas por sua equipe técnica) é juridicamente mais adequada que a proposta de decisão da IA que veio a ser descartada. Além disso, a IA judicial⁵⁶ cumpre função de apontamento de vieses do julgador⁵⁷ (o que é interesse da democracia) e de casos em que ocorre a fuga injustificada da *ratio decidendi* de precedentes, servindo como um passo justificatório que deve ser publicamente acessível, a fim de que a tentação à simples desconsideração da racionalidade (ou irracionalidade) proposta pela IA seja inibida.

O grau de acesso às informações deve ser pensado de modo que as partes e os advogados tenham automaticamente informações sobre o que precisam saber

⁵⁴ Hartmann Peixoto, Fabiano; Silva, Roberta Zumblick Martins da. Inteligência Artificial e Direito. Vol. 1. Curitiba: Alteridade Editora, 2019, p. 40.

⁵⁵ Deeks, Ashley. Op cit., p. 1833.

⁵⁶ Seja um sistema/modelo de IA mais ou menos complexo, neste caso o foco é na sua aplicabilidade para o aprimoramento da decisão judicial, pois a IA judicial deve ser construída em benefício da sociedade da correta aplicação do Direito, não apenas para “acelerar” processos. O respeito à razoável duração do processo mediante a aceleração dos atos processuais é apenas uma das facetas positivas da implementação de sistemas de IA judicial. Otimização do processo decisório é mais, muito mais, do que apenas julgar rapidamente.

⁵⁷ Sunstein, Cass. Op. cit.

acerca das principais variáveis que influenciam cada julgamento com IA⁵⁸, uma vez que caso haja a demanda por informações adicionais, estas devem ser disponibilizadas para se garantir a legitimidade⁵⁹ democrática inerente à transparência completa (*full transparency*), eis que mesmo que os destinatários da decisão não consigam compreender a forma de funcionamento de um sistema de IA judicial (um modelo baseado uma rede neural, por exemplo), há a percepção de que o Judiciário, instituição democraticamente legítima⁶⁰ para a função, respeitou ativamente a transparência em seus julgamentos.⁶¹ Em resumo, o Judiciário da era da IA não pode ter receio do escrutínio público, ainda que existam dificuldades de compreensão do funcionamento de modelos de inteligência artificial judicial e do próprio Direito.

Dessa forma, tendo em vista que o uso da IA judicial, caso esta seja implementado sem a devida observância do direito à explicabilidade, há um risco ao senso de justiça dos cidadãos⁶², o que pode abalar a confiança nas decisões judiciais.

Para além do cuidado em relação aos riscos, os quais devem ser catalogados, conhecidos e monitorados maneira ativa, a IA de auxílio ao julgamento pode incorporar a extensão legislação aplicável a determinada classe de problemas e, inclusive, determinados valores para dentro do modelo⁶³, seja numa abordagem *machine learning* (com treinamento em extensa base de dados) ou mediante representação do conhecimento no próprio algoritmo, pautando-se sempre na possibilidade de pluralidade democrática (art. 7º da Resolução 332/2020 do CNJ).

3 RAZÕES PÚBLICAS E INTELIGÊNCIA ARTIFICIAL JUDICIAL: A EXPLICABILIDADE COMO ELEMENTO DA JUSTIÇA

A necessidade do desenvolvimento de sistemas inteligência artificial explicável (xAI) para a modelos aplicados ao Judiciário se amolda à necessária construção de decisões judiciais mediante a exposição de razões públicas, sendo que

⁵⁸ Licht, K. de Fine; Licht, J. de Fine. Op. cit.

⁵⁹ Id.

⁶⁰ Lorenzetto, Bruno Meneses; Clève, Clèmerson Merlin. “Teorias Interpretativas, Capacidades Institucionais e Crítica”, em Revista Direitos Fundamentais & Democracia, v. 19, n. 19, 2016, Curitiba, p. 131-168.

⁶¹ Licht, K. de Fine; Licht, J. de Fine. Op. cit.

⁶² Deeks, Ashley. Op. cit., p. 1833.

⁶³ Id.

estas têm o potencial de trazer ao centro do debate democrático aspectos, desde os mais básicos, que influenciam nos posicionamentos dos julgadores.

Nesse caminho, a ideia de explicabilidade como elemento das razões públicas das decisões auxiliadas pela IA é uma possibilidade que se adequa à democracia constitucional, pois “a razão pública é a razão de cidadãos iguais que, enquanto corpo coletivo, exercem um poder político final e coercitivo uns sobre os outros ao promulgar leis e emendar sua constituição”.⁶⁴

A explicabilidade nasce como elemento intrínseco às razões públicas, na medida em que o Judiciário, órgão constitucional de quem se espera racionalidade⁶⁵ e sinceridade⁶⁶ inerente à boa-fé argumentativa, tem o ônus de explicar os motivos de suas decisões (art. 93, IX, e art. 489, *caput*, II, §1º e §2º, do CPC). Dessa maneira, devido aos aspectos sensíveis adicionais trazidos pela inclusão da IA judicial no processo decisório, é necessário o estímulo à construção das decisões mediante razões públicas que possam ser razoavelmente aceitas na sociedade.⁶⁷

O processo de decisão através de razões públicas leva à formação de consenso acerca das decisões judiciais, consenso este que será a fonte de inspiração para a programação do sistema de IA e/ou para os critérios de seleção e tratamento das bases de dados utilizadas no aprendizado de máquina. Tal consenso não é estático, mas sim dinâmico, uma vez que se trata de construção reiterada de um **consenso sobreposto**⁶⁸ numa democracia viva, onde há a sedimentação das estruturas jurídico-organizacionais do Estado e da sociedade de maneira paulatina, camada por camada.

Nessa linha, mediante a exposição consciente das razões das decisões tomadas com auxílio de inteligência artificial, através das razões públicas: **[a]** pelas quais determinado algoritmo (de representação de conhecimento ou de técnica de *machine learning*) foi adotado em detrimento de outras abordagens no código-fonte; **[b]** de por que determinada base de dados foi utilizada no treinamento, no teste ou na validação de modelo baseado em aprendizado de máquina; **[c]** das causas de

⁶⁴ Rawls, John. O Liberalismo Político. 2.ª ed. São Paulo: Editora Ática, 2000, p. 263.

⁶⁵ Alexy, Robert. Teoria da Argumentação Jurídica: A Teoria do Discurso Racional como Teoria da Justificação Jurídica. São Paulo: Landy, 2005, p. 281.

⁶⁶ *Ibid.*, p. 280.

⁶⁷ Rawls, John. O Liberalismo Político, p. 266-267.

⁶⁸ Gargarella, Roberto. *Op. cit.*, p. 236-238.

descontinuação ou não implementação de modelos de IA judicial; e, **[d]** na medida da possibilidade da técnica, porque determinada decisão foi proposta por modelo de IA judicial (resultado do treinamento, espécie de estruturação algorítmica adotada, base de dados utilizada, escolha do juiz-desenvolvedor, escolha centralizada pela administração do Tribunal, etc.).

Desse modo, a abordagem através da construção democrática de razões públicas tem potencial de otimizar o direito à explicabilidade, animando-o com exigência de conteúdo efetivamente democrático ao respeitar direitos fundamentais, conteúdo este que possa ser aceito como “razoável” por qualquer pessoa num debate público, bem como se caracterizando como um processo de construção das decisões que leve em conta a pluralidade. Nesse sentido, as razões públicas são uma garantia de que no processo decisão acerca da adoção da IA e nas decisões judiciais auxiliadas por IA haverá um ônus democrático que deverá ser cumprido pelos tribunais, em otimização à explicabilidade, o que tem o potencial de inibir abusos e a tentação de adoção de modelos de IA judicial voltados apenas para a produtividade numa perspectiva numérica, em detrimento do desenvolvimento de sistemas computacionais que tragam mais segurança numa arquitetura de escolhas que estimule a decisão correta (racional)⁶⁹ explicitamente justificada⁷⁰ e consciente pela pessoa do julgador.

Referido processo público de construção da explicabilidade dos modelos de IA judicial envolve **[a]** a política de tecnologia da informação de cada Tribunal; **[b]** o projeto do sistema de IA judicial; e **[c]** a implementação de técnicas de explicabilidade (*xAI*) para os resultados (*outputs*) de processamento de decisões via inteligência artificial.

Especificamente em relação à explicabilidade das decisões auxiliadas por IA, deve-se indicar a separação entre as razões públicas e as não públicas (de igrejas, associações da sociedade civil e entidades representativas de grupos socialmente atomizados⁷¹), pois as razões não públicas são vitais para a construção do debate, são parte da cultura de fundo que subsidia as visões de mundo a serem ouvidas com

⁶⁹ Alexy, Robert. Teoria da Argumentação Jurídica, p. 280.

⁷⁰ Bretas, Carla Panza; Mattos, Kennedy Josue Greca de; Schier, Paulo Ricardo. “Teoria do Direito e da Constituição no contexto do neoconstitucionalismo pós-positivista: análise quanto à racionalidade”, em Revista da Faculdade de Direito da UFRGS, v. 35, 2016, Porto Alegre, p. 115-130.

⁷¹ Rawls, John. O Liberalismo Político, p. 261-262.

base na alteridade social⁷², mas não são as razões aptas a serem adotadas numa sociedade democrática. Tal escrutínio no processo de desenvolvimento dos sistemas de IA judicial deve ser cauteloso, de maneira não discriminar minorias (art. 7º da Resolução 332/2020 do CNJ).

Ademais, quando se trata da explicabilidade de decisões judiciais específicas auxiliadas por IA judicial, a argumentação é o caminho adequado para a explicitação de quais razões *a priori* não públicas podem compor a construção da razão pública⁷³, pois existem elementos mínimos a caracterizar uma argumentação racional⁷⁴, afastando-se do mero discurso retórico político ou sectário.

A técnica argumentativa baseada em princípios da evidência (com critérios lógico-rationais) é parte da concepção política de Justiça⁷⁵, argumentação esta em que são admitidos valores da justiça política e valores da razão pública (o que inclui uma ideia de razoabilidade e o dever moral de civilidade)⁷⁶. Dessa forma, as diretrizes de indagação (princípios da argumentação e regras de evidência) são o elemento procedimental da concepção política de justiça “à luz dos quais os cidadãos devem julgar se os princípios substantivos aplicam-se de forma apropriada e identificar as leis e políticas que melhor os satisfaçam”.⁷⁷

Para além da concepção de posição original⁷⁸, mas aproveitando-se da exigência de racionalidade do momento crítico de escolha democrática das questões constitucionais básicas/fundamentais, mediante uma concepção política de justiça as partes devem agir com boa-fé argumentativa⁷⁹ de maneira que quem toma posições na sociedade com pretensão que uma decisão pública seja influenciada

“Isso significa que cada um de nós deve ter e deve estar preparado para explicar um critério acerca de que princípios e diretrizes pensamos que se pode razoavelmente esperar que os outros cidadãos (que também são livres e iguais) subscrevam junto conosco. Precisamos dispor de algum teste, que

⁷² Ibid., p. 264.

⁷³ Ibid., p. 269.

⁷⁴ Ibid., p. 269-270.

⁷⁵ Ibid. p. 273.

⁷⁶ Ibid. p. 273-274.

⁷⁷ Rawls, John. A Theory of Justice. Cambridge: Harvard University Press, 1971, p. 17-22.

⁷⁸ Gargarella, Roberto. Op. cit. p. 227.

⁷⁹ Rawls, John. O Liberalismo Político, p. 275.

estejamos dispostos a explicitar, para dizer quando essa condição é satisfeita⁸⁰.

Isso é necessário pois é inevitável, e desejável, que as pessoas tenham visões diferentes⁸¹, sendo um “debate ordenado entre elas ao longo do tempo é uma forma confiável de descobrir qual é a mais razoável, se alguma o é”.⁸²

Embora os cidadãos e o legislativo tenham maior margem de debate público, o Judiciário é criação de uma razão pública, sendo por ela limitado, o que o incumbe de grande ônus argumentativo, devendo o julgador se guiar exclusivamente pela razão pública⁸³

É papel dos Juízes, de forma refletida, buscar as melhores interpretações da legislação que puderem fazer⁸⁴ na concepção pública de justiça, pois o “papel do tribunal aqui é parte da publicidade da razão, e um aspecto do papel amplo ou educativo da razão pública”.⁸⁵ Nessa função constitucional os julgadores não podem se utilizar de sua moralidade particular⁸⁶, pois mesmo que venham a ocorrer dificuldades interpretativas deve-se resistir ao abandono da razão pública nos julgamentos.

Portanto, os sistemas de IA para auxílio ao julgamento devem ser pensados a partir da premissa de que o consenso do processo (numa dialética limitada e controlada a cada caso topicamente considerado) não supera o consenso contraposto formado de maneira objetiva na sociedade. Dessa maneira, o modelo de supervisão humana pelo julgador do caso (art. 19, Parágrafo único, da Resolução 332/2020 do CNJ) trará o equilíbrio necessário entre ambas as abordagens de desenvolvimento de sistemas de inteligência artificial judicial.

CONCLUSÃO

O presente estudo demonstrou os pontos de ligação entre a teoria da Justiça, na perspectiva de John Rawls em sua segunda fase de pensamento a partir da obra

⁸⁰ Ibid., p. 276.

⁸¹ Ibid., p. 277.

⁸² Id.

⁸³ Ibid. p. 286.

⁸⁴ Ibid. p. 286-287.

⁸⁵ Ibid., p. 287.

⁸⁶ Id.

O Liberalismo Político, e o direito à explicabilidade das decisões judiciais auxiliadas por inteligência artificial.

Foi observado que a explicabilidade das decisões não é problema totalmente novo, trata-se de questão tradicional acerca da transparência e da sinceridade na construção das decisões judiciais.

O direito à explicabilidade das decisões judiciais tomadas com auxílio de inteligência artificial é fomentado numa concepção de justiça política, a qual pressupõe a construção do consenso contraposto através de razões públicas, mediante a definição de um espaço democrático de argumentação marcada pela pluralidade.

O que ficou claro é que o fenômeno da regulação da inteligência artificial judicial, ainda que esteja experimentando recente abordagem regulatória (Conselho Nacional de Justiça, Lei Geral de Proteção de Dados, União Europeia, OCDE, UNESCO, etc.), demanda estudo típico da filosofia do Direito, a fim de serem sedimentadas as bases para o reconhecimento do direito fundamental à transparência das decisões tomadas com base em inteligência artificial.

Evidenciou-se que a nova camada de linguagem da decisão judicial, implementada através da inteligência artificial, traz consigo riscos (opacidade e vieses algorítmicos, por exemplo), os quais devem ser tratados na computação aplicada ao direito mediante soluções híbridas de mútua cooperação científica entre teoria do Direito e Ciência da Computação, o que é possível mediante arquiteturas de sistemas voltados para a explicabilidade de decisões tomadas por IA.

Nesse aspecto, a teoria das razões públicas aplicada à decisão judicial tem elementos que podem ser utilizados para potencializar as técnicas já existentes de inteligência artificial explicável (xAI), desde o processo de desenvolvimento dos algoritmos e modelos de IA até os resultados de processamento computacional.

Dessa forma, uma IA judicial pode empregar abordagens procedimental, de representação do conhecimento ou de *machine learning* (incluindo-se o dilúvio de dados necessário aos modelos *deep learning*) em relação à sua programação, todavia deve-se sempre desenvolver os sistemas com a possibilidade de explicabilidade compreensível aos humanos.

Neste ponto, até mesmo para que adequadamente existam desenvolvimento, controle e supervisão de sistemas de IA judicial, é caminho natural que os profissionais

do Direito do futuro (e do presente) passem da figura do operador do direito para posições mais próximas do que se pode chamar de desenvolvedor de software jurídico.

A crítica ao entrelaçamento entre direito e inteligência artificial continuará a ter sua relevância, mas, se ficar postada em posição exclusivamente reacionária, em pouco tempo será vencida pela realidade tecnológica que se impõe.

Portanto, a IA de auxílio à decisão judicial (pensamento/processamento híbrido humano-máquina) já funciona com as limitações que a tecnologia impõe, cabendo desde já reconhecer a explicabilidade da IA jurídica como um direito necessário para uma sociedade plural e justa.

REFERÊNCIAS

Alexy, Robert. Teoria da Argumentação Jurídica: A Teoria do Discurso Racional como Teoria da Justificação Jurídica. São Paulo: Landy, 2005.

_____. Teoria dos Direitos Fundamentais. 2ª ed. São Paulo, 2011.

Ashley, Kevin. Artificial Intelligence and Legal Analytics: New tools for law practice in digital age. Cambridge: Cambridge University Press, 2017.

Bostrom, Nick. Superinteligência: Caminhos, perigos, estratégias. Rio de Janeiro: Darkside Books, 2014.

Brasil. Conselho Nacional de Justiça. Resolução 332, de 21 de agosto de 2020. Dispõe sobre a ética, a transparência e a governança na produção e no uso de Inteligência Artificial no Poder Judiciário e dá outras providências.

Bretas, Carla Panza; Mattos, Kennedy Josue Greca de; Schier, Paulo Ricardo. “Teoria do Direito e da Constituição no contexto do neoconstitucionalismo pós-positivista: análise quanto à racionalidade”, em Revista da Faculdade de Direito da UFRGS, v. 35, Porto Alegre, dez. 2016, p. 115-130.

Brown, Tom B. et al. “Language Models are Few-Shot Learners”, OpenAI, San Francisco, arXiv:2005.14165. Disponível em: <https://arxiv.org/pdf/2005.14165.pdf>. Acesso em 05/04/2021.

Casetext. Cara A.I. Disponível em: <https://casetext.com/cara-ai>. Acesso em 28/03/2021.

Comissão Europeia. “Ethics Guidelines for Trustworthy AI”. 08 April 2019. Disponível em: https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419. Acesso em: 25/07/2020.

Deeks, Ashley. “The Judicial Demand for Explainable Artificial Intelligence”, em *Columbia Law Review*, v. 119, 2020, n. 7.

Dewey, Daniel; Russel, Stuart; Tegmark, Max. “Research Priorities for Robust and Beneficial Artificial Intelligence” em *AI Magazine*, v. 36, n. 4, 2015.

Future of Life Institute. “Asilomar AI Principles”. Disponível em: <https://futureoflife.org/ai-principles>. Acesso em: 05/08/2020.

Gargarella, Roberto. *As Teorias da Justiça Depois de Rawls: Um breve manual de filosofia política*. São Paulo: WMF Martins Fontes, 2008.

Gilpin, Leilani H. et. al. “Explaining Explanations: An Overview of Interpretability of Machine Learning”, em 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA), 2018, Turin, p. 80-89.

Habermas, Jürgen. *Teoria do Agir Comunicativo: Racionalidade da Ação e Racionalidade Social*. Vol. 1. São Paulo: WMF Martins Fontes, 2012.

Harari, Yuval Noah. *Homo Deus: Uma breve história do amanhã*. São Paulo: Companhia das Letras, 2016.

Hartmann Peixoto, Fabiano; Silva, Roberta Zumblick Martins da. *Inteligência Artificial e Direito*. Vol. 1. Curitiba: Alteridade Editora, 2019.

Howard, Philip N. *Pax Technica: How the Internet of Things May Set Us Free or Lock Us Up*. New Haven: Yale University Press, 2015.

Kahneman, Daniel. *Rápido e Devagar: Duas formas de pensar*. Tradução: Cássio de Arantes Leite. Rio de Janeiro: Objetiva, 2012.

Kurzweil, Ray. *A Singularidade Está Próxima: quando os humanos transcendem a biologia*. Tradução: Ana Goldberger. São Paulo: Iluminuras, 2018.

Lage, Fernanda de Carvalho. *Manual de Inteligência Artificial no Direito Brasileiro*. Salvador: Editora Juspodivm, 2021.

Licht, K. de Fine; Licht, J. de Fine. “Artificial intelligence, transparency, and public decision-making”, em *AI & Society*, Mar. 2020.

Lorenzetto, Bruno Meneses; Clève, Clèmerson Merlin. “Teorias Interpretativas, Capacidades Institucionais e Crítica”, em *Revista Direitos Fundamentais & Democracia*, v. 19, n. 19, 2016, Curitiba, p. 131-168.

MIT Technology Review. Por que a GPT-3 é o melhor e o pior da IA atualmente. 01/04/2021. Disponível em: <https://mittechreview.com.br/por-que-a-gpt-3-e-o-melhor-e-o-pior-da-ia-atualmente>. Acesso em 08/04/2021.

Newton, Isaac. "Letter from Sir Isaac Newton to Robert Hooke", 1675. Disponível em: <https://digitallibrary.hsp.org/index.php/Detail/objects/9792>. Acesso em: 06/04/2021.

OCDE. Recommendation 0449, of 21 May 2019: Recommendation of the Council on Artificial Intelligence. Disponível em: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>. Acesso em: 04/08/2020.

OpenAI. OpenAI API. Disponível em: <https://openai.com/blog/openai-api/>. Acesso em 08/04/2021.

Pasquale, Frank. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge: Harvard University Press, 2015.

Rawls, John. *A Theory of Justice*. Cambridge: Harvard University Press, 1971.

_____. *Justiça e Democracia*. São Paulo: Martins Fontes, 2000.

_____. *O Liberalismo Político*. 2.^a ed. São Paulo: Editora Ática, 2000.

Rover, Aires José. *Informática no Direito: Inteligência Artificial*. Curitiba: Juruá, 2001.

_____. "Machine Learning no Poder Judiciário: uma biblioteca temática, ano 2020", em Hartmann Peixoto, Fabiano (Org.). *Inteligência Artificial: estudos de inteligência artificial*. Vol. 4. Curitiba: Alteridade, 2021.

Russell, Stuart; NORVIG, Peter. *Inteligência Artificial*. 3^a ed. Rio de Janeiro: Elsevier, 2013.

Salomão, Luis Felipe (coord.) et al. *Inteligência Artificial: Tecnologia Aplicada à Gestão dos Conflitos no Âmbito do Poder Judiciário Brasileiro*. FGV Conhecimento: Centro de Inovação, Administração e Pesquisa do Judiciário, 2020.

Schwab, Klaus. *A Quarta Revolução Industrial*. São Paulo: Edipro, 2018.

SUNSTEIN, Cass R. "Algorithms, Correcting Biases", em *Social Research: Algorithms*, v. 86, n. 2, New York, 2019.

Susskind, Richard. *Online Courts and the Future of Justice*. Oxford: Oxford University Press.

Taulli, Tom. *Introdução à Inteligência Artificial: Uma abordagem não técnica*. São Paulo: Novatec, 2020.

UNESCO. "First version of a draft text of a recommendation on the Ethics of Artificial Intelligence". 15 May 2020. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000373434>. Acesso em: 25/07/2020.

Wagner, Ben. "Accountability by design in technology research", em *Computer Law & Security Review*, v. 37, 105398, ISSN 0267-3649, 2020, Oxford.

Yapo, Adrienne; Weiss, Joseph. "Ethical Implications of Bias in Machine Learning", em *Proceedings of 51st Hawaii International Conference on System Sciences*, Waikoloa Village, 2018, p. 5365-5372.