A INTELIGÊNCIA ARTIFICIAL E A CRISE DA DECISÃO JURÍDICA ÉTICA

ARTIFICIAL INTELLIGENCE AND THE CRISIS OF ETHICAL LEGAL DECISION-MAKING

Edielma Rocha Dutra[1]

Resumo: Este artigo investiga as transformações introduzidas pela Inteligência Artificial (IA) no âmbito da decisão jurídica e os desafios ético-filosóficos que emergem dessa interface. Parte-se da hipótese de que a automação crescente da decisão judicial coloca em crise a tradicional noção de julgamento humano, com suas bases axiológicas, morais e de responsabilidade. O objetivo do estudo é analisar os impactos éticos e filosóficos decorrentes da utilização da IA no processo decisório jurídico contemporâneo, identificando as principais problemáticas e propondo diretrizes que contribuam para mitigar a crise ética que se instaura. No desenvolvimento, examina-se o fenômeno da delegação decisória a algoritmos, a perda de transparência e explicabilidade, bem como os riscos de viés e desresponsabilização. A metodologia adotada é de natureza bibliográfica e analítica, com abordagem qualitativa e enfoque interdisciplinar entre Direito, Filosofia e Ética. Conclui-se que, para preservar a integridade ética da decisão jurídica, faz-se necessário afirmar a supervisão humana, a explicabilidade dos sistemas de IA e a construção de marcos regulatórios eficientes que garantam os valores da dignidade humana, do devido processo legal e da justiça.

Palavras-Chave: Inteligência Artificial; decisão jurídica; ética; responsabilidade; justiça.

Abstract: This article investigates the transformations introduced by Artificial Intelligence (AI) in the field of legal decision-making and the ethical-philosophical challenges that emerge from this interface. It is based on the hypothesis that the growing automation of judicial decision-making puts into crisis the traditional notion of human judgment, with its axiological, moral, and responsibility-based foundations. The objective of the study is to analyze the ethical and philosophical impacts resulting from the use of AI in contemporary legal decision-making processes, identifying the main issues and proposing guidelines to mitigate the ensuing ethical crisis. The research examines the phenomenon of decision-making delegation to algorithms, the loss of transparency and explainability, as well as the risks of bias and lack of accountability. The adopted methodology is bibliographic and analytical in nature, with a qualitative approach and an interdisciplinary focus encompassing Law, Philosophy, and Ethics. It concludes that, in order to preserve the ethical integrity of legal decision-making, it is necessary to ensure human supervision, the explainability of AI systems, and the establishment of efficient regulatory frameworks that safeguard the values of human dignity, due process of law, and justice.

Keywords: Artificial Intelligence; legal decision-making; ethics; responsibility; justice.

INTRODUÇÃO

A inteligência artificial (IA) tem se consolidado como uma das mais impactantes transformações tecnológicas do século XXI, influenciando não apenas os campos

[1] Graduada em Direito pela Universidade Católica de Santos – UNISANTOS. Especialista em Direito Processual Penal pelo Instituto Damásio de Direito/IBMEC. Especialista em Direito Penal e Criminologia pela Pontifícia Universidade Católica do Rio Grande do Sul – PUC/RS. Mestranda em Direito da Sociedade da Informação nas Faculdades Metropolitanas Unidas – FMU. Pesquisadora.

científicos e econômicos, mas também os sistemas jurídicos e éticos que sustentam a vida em sociedade.

No contexto do Direito, o uso crescente de algoritmos e sistemas automatizados na análise, predição e, em alguns casos, na própria formulação de decisões judiciais, suscita um debate profundo acerca da natureza do ato de julgar e da legitimidade da decisão jurídica em um cenário de automatização. A máquina, ao reproduzir padrões e probabilidades extraídas de dados, introduz uma lógica de cálculo e eficiência que pode colidir com os fundamentos axiológicos e morais da decisão humana. A decisão jurídica, compreendida como um ato de vontade racional e moralmente orientado, é o resultado de uma ponderação entre normas, fatos e valores, na qual o julgador exerce não apenas uma função técnica, mas também ética. Essa dimensão ética é o que confere humanidade, prudência e justiça à decisão.

Quando a inteligência artificial passa a intermediar ou influenciar esse processo, instala-se uma tensão entre o juízo humano, pautado por princípios, empatia e interpretação, e o juízo algorítmico, regido por padrões de eficiência e lógica probabilística. A introdução da IA no âmbito da jurisdição não representa apenas um avanço tecnológico, mas uma inflexão paradigmática, que desafia a própria compreensão do que significa decidir de forma justa.

A chamada "crise da decisão jurídica ética" manifesta-se, portanto, na medida em que os sistemas de IA, ainda que concebidos como instrumentos de apoio, começam a interferir no espaço tradicionalmente reservado à consciência e à responsabilidade humanas. A decisão judicial, que sempre foi expressão da racionalidade prática (vinculada à noção aristotélica de *phronesis*), passa a ser tensionada por uma racionalidade instrumental, típica das máquinas. Essa substituição simbólica do discernimento humano pelo cálculo técnico questiona se é possível falar em justiça quando o processo decisório se torna opaco, automatizado e, muitas vezes, ininteligível para seus destinatários. Além do aspecto filosófico, há uma dimensão ética e jurídica concreta nesse fenômeno. A opacidade dos algoritmos, conhecida como *black box problem*, dificulta a compreensão e a contestação das decisões automatizadas, afrontando princípios como a motivação das decisões judiciais, a transparência e o devido processo legal.

Ademais, a dependência crescente de dados históricos pode perpetuar desigualdades e vieses estruturais, uma vez que os algoritmos aprendem a partir de padrões sociais já marcados por assimetrias de poder e discriminação. O risco de reproduzir injustiças sob a aparência de neutralidade tecnológica é real e exige reflexão crítica sobre o papel da IA no sistema jurídico. Neste cenário, impõe-se uma análise interdisciplinar que articule Direito, Filosofia e Ética, de modo a compreender a

profundidade da crise instaurada pela automação da decisão jurídica.

O debate não se resume à eficiência técnica dos sistemas, mas à redefinição do próprio conceito de justiça e de responsabilidade. Se a decisão jurídica é um ato ético, fundado na capacidade humana de deliberar e justificar, a substituição ou o enfraquecimento desse elemento por uma racionalidade algorítmica põe em risco a legitimidade do sistema jurídico como espaço de realização dos direitos fundamentais e da dignidade humana.

A presente pesquisa parte da hipótese de que a incorporação da inteligência artificial no processo de decisão jurídica representa uma crise ética e filosófica de grande relevância, vez que tensiona os fundamentos normativos e morais que sustentam a função jurisdicional. Propõe-se examinar, sob um viés crítico e interdisciplinar, como a automação da decisão afeta os princípios da transparência, responsabilidade e justiça, e refletir sobre possíveis caminhos regulatórios e éticos que permitam a integração responsável da IA ao Direito.

Trata-se, portanto, de investigar não apenas os efeitos técnicos da tecnologia, mas, sobretudo, o modo como ela redefine a própria ideia de decisão jurídica e a essência do ato de julgar no contexto do Direito contemporâneo.

1 FUNDAMENTOS CONCEITUAIS

A compreensão da crise ética na decisão jurídica mediada por inteligência artificial exige o delineamento conceitual dos elementos que a compõem: a natureza da IA e seu funcionamento na esfera jurídica, a estrutura ética do ato de julgar e a forma pela qual essas duas dimensões entram em tensão no contexto contemporâneo.

O debate em torno da automação decisória não se restringe à aplicação técnica da IA, mas envolve uma questão ontológica e axiológica — o que significa decidir e quem pode, legitimamente, fazê-lo dentro de um sistema jurídico orientado por valores humanos.

A IA, em sentido amplo, designa sistemas computacionais capazes de realizar tarefas que exigiriam, em condições ordinárias, algum grau de inteligência humana, tais como raciocínio, aprendizado, interpretação de padrões e tomada de decisão. Explica Barreto Junior:

Inteligência artificial é a nomenclatura pela qual se convencionou chamar a tecnologia informática desenvolvida com o intuito de oferecer soluções para perguntas humanas, com crescente probabilidade estatística de acerto, questões cujas respostas exigem a simulação da capacidade humana de raciocinar, perceber, tomar decisões e resolver problemas. A nomenclatura *Inteligência Artificial* não é propriamente nova.

Originalmente foi aplicada a processos de automação e robótica desenvolvidos desde a segunda metade do século XX (Barreto Junior, 2024, p. 83).

A IA pode ser definida, portanto, com "um conjunto de artefatos projetados para reproduzir ou ampliar capacidades cognitivas humanas, mediante processos algorítmicos e modelos de aprendizado de máquina. Essa definição destaca o caráter instrumental da IA: ela não pensa nem delibera, mas opera por meio da análise de dados e da inferência estatística, convertendo informações em previsões e classificações. Ademais, justificar razões é orientar condutas, inserindo-se na racionalidade prática, em que a argumentação jurídica, conforme Perelman, vai além da lógica formal e aplica a razão a casos concretos:

[...] admitir a possibilidade de uma justificação racional significa admitir ao mesmo tempo um uso pratico da razão, não limitando esta à faculdade de discernir relações necessárias, nem sequer relações referentes ao verdadeiro ou ao falso. Isso porque toda justificação racional supõe que raciocinar não e somente demonstrar e calcular, e também deliberar, criar e refutar, e apresentar razoes pro e contra, e, numa palavra, argumentar. A ideia de justificação racional e, de fato, inseparável da ideia de argumentação racional (Perelman, 1996, p. 344).

No campo jurídico, a aplicação da IA tem avançado de forma significativa. Ferramentas de análise preditiva, triagem processual e recomendação de sentenças são utilizadas em diversos países, inclusive no Brasil, em projetos como o Victor, desenvolvido pelo Supremo Tribunal Federal (STF), e o Sinapses, do Conselho Nacional de Justiça (CNJ). Tais sistemas têm o mérito de otimizar fluxos e reduzir a sobrecarga de processos, mas levantam questionamentos quanto à preservação da autonomia decisória do julgador (Valle; Fuentes i Gasó; Ajus, 2023, pp. 12 e 21). Siqueira, Morais e Santos refletem acerca do devido processo legal e o julgamento delegado à "máquina":

A vinculação primeira do princípio do devido processo legal como fundamento e limite ao exercício da tomada de decisão judicial lastreada em IA é impositiva, face ao fenômeno da filtragem constitucional, especialmente considerando a necessidade de se ponderar até que ponto máquinas podem substituir humanos no ato de julgar em um sistema jurídico em que a pessoa é ponto de partida e de chegada das ações do Estado (Siqueira, Morais e Santos, 2022, p. 10).

Os autores também advertem que a inteligência artificial não pode substituir o dever analítico de fundamentação da decisão judicial, sob pena de suprimir o elemento ético da jurisdição e desfigurar o papel humano de interpretação e valoração. Conforme explicam:

Não se pode descuidar, contudo, que a utilização exponencial de dados e procedimentos automatizados não pode afastar-se, assim, da participação humana. O

devido processo legal é a mola propulsora do dever de transparência, correção e legitimidade dos procedimentos que versam sobre IA. As soluções e sugestões decorrentes do uso da IA decorrem de algoritmos, ou seja, de um conjunto de instruções, projeções, indicações, sequência de regras alimentadas pelo ser humano. O que pode conter vieses cognitivos. Assim, os vieses podem naturalmente afetar as soluções apresentadas pela IA. Circunstância que redobra a necessidade de cautela no uso de IA na tomada de decisão judicial e reacende a preocupação de reservá-la para a tomada de decisões repetitivas destituídas de complexidade, sem prejuízo das restrições apontadas no item precedente envolvendo fundamentação analítica (Siqueira; Morais; Santos, 2022, p. 24).

Do ponto de vista filosófico, a decisão jurídica é compreendida como uma prática de racionalidade prática, no sentido aristotélico da *phronesis* (sabedoria deliberativa orientada pelo bem e pela justiça). O núcleo da filosofia prática de Aristóteles está no elemento denominado como *phronesis,* o elemento concebido como uma forma de racionalidade prática orientada para a ação. Aristóteles observa que "chegaremos à definição de *phronesis* se considerarmos quais são as pessoas dotadas dessa forma de virtude" (Aristóteles, 2018, p. 116). Aristóteles, inclusive, aprofunda a questão da deliberação (julgamento) na presença do íntimo humano:

Pensa-se que é característico de uma pessoa dotada de *phronesis* ser capaz de deliberar bem acerca do que é bom e conveniente para si mesma, não em relação a um aspecto particular – por exemplo, quando se quer saber quais as espécies de coisas que concorrem para a saúde e para o vigor físico –, e sim acerca das espécies de coisas que nos levam a viver bem de modo geral (Aristóteles, 2018, p. 116).

O julgador, ao decidir, não apenas aplica normas, mas pondera princípios e valores, construindo sentido jurídico a partir do caso concreto. Dworkin descreve essa função como um exercício de integridade moral e interpretativa, no qual o juiz busca a melhor resposta possível na coerência do sistema jurídico. Julgar é um ato ético que exige do juiz responsabilidade moral para equilibrar justiça, equidade e devido processo, garantindo decisões corretas e legítimas:

A justiça diz respeito ao resultado correto do sistema político: a distribuição correta de bens, oportunidades e outros recursos. A equidade é uma questão da estrutura correta para esse sistema, a estrutura que distribui a influência sobre as decisões políticas da maneira adequada. O devido processo legal adjetivo é uma questão dos procedimentos corretos para a aplicação de regras e regulamentos que o sistema produziu. A supremacia legislativa que obriga Hércules a aplicar as leis, mesmo quando produz uma incoerência substantiva, é uma questão de equidade porque protege o poder da maioria de fazer o direito que quer. As doutrinas rigorosas do precedente, as práticas da história legislativa e

a prioridade local são em grande parte, embora de maneira distintas, questões de processo legal adjetivo, porque estimulam os cidadãos a confiar em suposições e pronunciamentos doutrinários que seria errado trair ao julgá-los depois do fato (Dworkin, 2021, p. 483).

A ética, por sua vez, desempenha papel fundante no exercício jurisdicional. Habermas ressalta que toda decisão legítima deve ser comunicativamente justificável, pois, como a ética moderna exige uma apropriação consciente e autocrítica da própria vida, o ato de decidir no âmbito jurídico requer uma fundamentação racional e responsável diante da comunidade:

What was considered "ethics" since the time of Aristotle now assumed a new, subjectivistic sense. This was true of both individual life histories and of intersubjectively shared traditions and forms of life. In connection with, and in reaction to, a growing autobiographical literature of confessions and self-examinations—running from Rousseau through Kierkegaard to Sartre—a kind of reflection developed that altered attitudes toward one's own life. To put it briefly, in place of exemplary instructions in the virtuous life and recommended models of the good life, one finds an increasingly pronounced, abstract demand for a conscious, self-critical appropriation, the demand that one responsibly take possession of one's own individual, irreplaceable, and contingent life history. Radicalized interiority is burdened with the task of achieving a self-understanding in which self-knowledge and existential decision interpenetrate (Habermas, 1996, p 96).[2]

A decisão jurídica ética é aquela que preserva a transparência, a possibilidade de revisão e a abertura ao diálogo racional. Em contrapartida, a decisão algorítmica, ainda que eficiente, tende a se tornar opaca, automatizada e desprovida de justificação compreensível. Essa opacidade, denominada *black box effect*, desafia diretamente a noção de responsabilidade e de *accountability*. A questão é que o ato de julgar, raciocinar, fazer relações, vai além do simples fato de raciocinar. Por vezes, é necessário utilizar os cinco sentidos: visão, tato, paladar, olfato, audição. Este ato é complexo. Sousa corrobora este pensamento:

De facto, há muito a neurociência procura mapear o processo humano de tomada de decisão, sendo tarefa difícil, não obstante os avanços já alcançados. Os sistemas mais

[2] Tradução livre: O que se considerava “ética” desde a época de Aristóteles passou a assumir um novo sentido, de caráter mais subjetivista. Isso se aplicava tanto às histórias de vida individuais quanto às tradições e formas de vida compartilhadas intersubjetivamente. Em conexão com — e em reação a — uma crescente literatura autobiográfica de confissões e autoexames, que vai de Rousseau a Kierkegaard e Sartre, desenvolveu-se um tipo de reflexão que modificou as atitudes em relação à própria vida. Em resumo, no lugar das instruções exemplares sobre a vida virtuosa e dos modelos recomendados de boa conduta, surgiu uma exigência cada vez mais abstrata e enfática de apropriação consciente e autocrítica: a exigência de que cada um assuma, de forma responsável, a posse de sua própria história de vida — individual, insubstituível e contingente. Essa interioridade radicalizada passou a carregar o encargo de alcançar uma autocompreensão em que o autoconhecimento e a decisão existencial se interpenetram

complexos de inteligência artificial convocam um problema semelhante na medida em que não é possível mapear o processo que levou à elaboração do resultado. Por exemplo, a *deep neuronal network* integra uma arquitetura complexa e multidimensional de "neurónios artificiais", capazes de aprender com a experiência, interligando e relacionando incontável informação (Sousa, 2020, p. 68).

Não só isso. O avanço tecnológico intensificou a complexidade e a opacidade dos sistemas modernos, tornando mais difícil compreender os mecanismos que produzem determinados resultados e desafiando os critérios tradicionais como, por exemplo, o da causalidade no Direito Penal. Sousa avalia esta relação pontuando que sistemas, ainda em fase relativamente recente de consolidação, foram significativamente impulsionados pelo avanço tecnológico. Contudo, à medida que se tornam mais complexos, opacos e autônomos, torna-se igualmente mais difícil identificar de maneira precisa o processo que conduziu ao resultado produzido. Tal dificuldade vem sendo apontada por parte da doutrina como um entrave à própria verificação do nexo causal segundo os critérios tradicionalmente adotados no direito penal. Esses questionamentos se tornam ainda mais relevantes quando analisados sob perspectivas normativo-causais que pressupõem a previsibilidade do resultado, isto é, quando o estabelecimento da causalidade depende da regularidade dos fatos e das regras de experiência — como exemplifica a teoria da causalidade adequada (Sousa, 2020, p. 68).

Surge, então, a denominada "crise da decisão jurídica ética". Essa crise pode ser entendida como o enfraquecimento do papel moral e racional do julgador diante da ascensão de uma racionalidade técnico-instrumental, na qual o valor da decisão é medido por sua eficiência e previsibilidade, e não por sua justiça. A substituição parcial do discernimento humano pelo cálculo algorítmico ameaça princípios estruturantes do processo judicial, como a motivação das decisões (art. 93, IX, da Constituição Federal de 1988) e o devido processo legal (art. 5º, LIV e LV). O risco ético consiste em reduzir o julgamento à mera função de confirmação estatística, destituindo-o de reflexão moral e crítica (Brasil, 2025).

Para Zanon, Branco e Silva, "não seria ousado referir que não há decisão tomada por uma máquina que seja totalmente neutra, ou seja, não vinculada à experiência do julgador, visto que os dados que sustentam as IAs são resultantes de interpretações humanas" (Zanon; Branco; Silva, 2025, p. 189). Os vieses morais e cognitivos incorporados aos algoritmos de IA revelam que a pretensa neutralidade tecnológica é uma ilusão metodológica: toda decisão automatizada é, em alguma medida, fruto de escolhas humanas e valores embutidos no código. Essa constatação reforça a necessidade de compreender a IA não como sujeito de decisão, mas como instrumento de apoio, cuja

utilização deve ser permanentemente acompanhada de controle ético, explicabilidade técnica e supervisão humana.

Nesse sentido, autores como Teigão e Fogaça (2025) defendem que a ética da decisão jurídica automatizada deve se basear em três pilares: (a) transparência algorítmica e auditabilidade, (b) presença de supervisão humana significativa e (c) preservação dos direitos fundamentais dos jurisdicionados. Tais parâmetros buscam reconduzir a tecnologia ao serviço da justiça, e não o contrário. A decisão jurídica ética, mesmo em um contexto de inteligência artificial, continua a depender de valores humanos, responsabilidade e reflexão moral. Ou seja, elementos insubstituíveis pela máquina (Teigão; Fogaça, 2025).

Portanto, compreender os fundamentos conceituais da IA e da ética decisória é essencial para identificar limites e possibilidades de sua interação no Direito contemporâneo. A crise da decisão jurídica ética não é apenas um problema técnico, mas um desafio epistemológico e moral. Exige repensar o papel do humano na era da decisão automatizada.

2 DESAFIOS ÉTICOS DA INTELIGÊNCIA ARTIFICIAL NA DECISÃO JURÍDICA

A aplicação da Inteligência Artificial (IA) no campo jurídico impõe uma reflexão ética profunda sobre os limites da tecnologia no ato de julgar. A progressiva delegação de funções decisórias a sistemas automatizados desafia princípios estruturantes do Direito, como a responsabilidade, a dignidade humana e a imparcialidade judicial. Tal fenômeno representa o confronto entre a racionalidade técnica, pautada pela eficiência e pela previsibilidade, e a racionalidade prática, orientada pelos valores morais e pelo discernimento ético do julgador. Vivencia-se a era da "hipernormatização artificial". Conforme explica Morais:

A hipernormatização artificial é o fenômeno que ocorre quando os elementos fáticos ou normativos considerados numa decisão jurídica não são adequadamente formalizados em uma premissa para o uso por um programa de Inteligência Artificial - IA. Por conta disso, a IA passa a operar com uma premissa artificial que é progressivamente hiperestabilizada diante do funcionamento eficaz do programa (Morais, 2020, pp. 51-52).

Surge a questão sobre o risco de uma hipernormatização artificial, que pode se manifestar por elementos normativos ou fáticos, embora seja inegável que essas IAs apresentem alta eficiência. Os elementos normativos aparecem quando se estabelece uma regra que passou por alterações interpretativas pelo tribunal ou sobre a qual não há

consenso acerca da interpretação adotada (Morais, 2020, p. 52). A primeira questão ética que emerge refere-se à opacidade dos algoritmos, fenômeno conhecido como *black box problem*. Trata-se da dificuldade (ou impossibilidade) de compreender de modo transparente os critérios internos utilizados pelos sistemas de aprendizado de máquina para produzir resultados ou recomendações. A limitação da autonomia das máquinas e a necessidade de reversibilidade das decisões reforçam que a opacidade algorítmica, o que compromete a legitimidade das decisões automatizadas ao impedir a verificação de seus fundamentos. Segundo Floridi e Cowls:

[...] the autonomy of humans should be promoted and that the autonomy of machines should be restricted and made intrinsically reversible, should human autonomy need to be protected or re-established (consider the case of a pilot able to turn off the automatic pilot and regain full control of the airplane). This introduces a notion we might call 'meta-autonomy,' or a 'decide-to-delegate' model: humans should retain the power to decide which decisions to take: exercising the freedom to choose where necessary, and ceding it in cases where overriding reasons, such as efficacy, may outweigh the loss of control over decision-making. Any delegation should also remain overridable in principle (i.e., deciding to decide again) (Floridi; Cowls, 2019).[3]

No contexto jurídico, tal opacidade desafia diretamente o dever constitucional de motivação das decisões, previsto no art. 93, IX, da Constituição Federal (Brasil, 2025). Se o julgador não é capaz de compreender plenamente as bases da resposta fornecida por um sistema de IA, a decisão resultante carece de fundamentação racional e, portanto, de legitimidade ética e jurídica. Além da opacidade, destaca-se o problema dos vieses algorítmicos, ou *algorithmic bias*. Como observa Zuboff, na obra "Capitalismo de Vigilância" (2021), os algoritmos "aprendem" a partir de dados históricos, refletindo inevitavelmente as desigualdades e preconceitos existentes na sociedade. No âmbito judicial, essa reprodução inconsciente de vieses pode conduzir à discriminação automatizada, afetando a isonomia e o direito à igualdade.

Estudos empíricos nos Estados Unidos demonstraram que softwares de previsão de reincidência, como o sistema *COMPAS*, atribuíram maior probabilidade de reincidência

[3] Tradução livre: a autonomia humana deve ser promovida, enquanto a autonomia das máquinas deve ser limitada e intrinsecamente reversível, caso seja necessário proteger ou restaurar a autonomia humana (como no caso de um piloto que pode desligar o piloto automático e reassumir o controle total do avião). Isso introduz uma noção que poderíamos chamar de "meta-autonomia", ou um modelo de "decidir-para-delegar": os humanos devem manter o poder de decidir quais decisões tomar, exercendo a liberdade de escolha quando necessário e cedendo-a em situações nas quais razões predominantes, como a eficácia, possam superar a perda de controle sobre a tomada de decisão. Qualquer delegação também deve permanecer, em princípio, passível de reversão (isto é, a possibilidade de decidir novamente).

a réus negros em comparação a réus brancos, mesmo diante de antecedentes semelhantes. Tal exemplo revela como a suposta neutralidade tecnológica pode mascarar desigualdades estruturais, transformando a IA em um instrumento de injustiça automatizada (Angwin; Larson; Mattu; Kirchner, 2016). Mesmo sistemas que utilizam, como fonte de consulta, um amplo conjunto de decisões anteriores proferidas por diversos magistrados não estão isentos da possibilidade de reproduzir vieses. Um exemplo de como um algoritmo baseado nesse tipo de acervo — e em outros critérios — pode apresentar viés discriminatório é o caso do *Correctional Offender Management Profiling for Alternative Sanctions* (COMPAS), empregado nos Estados Unidos. Esse software tem por finalidade avaliar o risco de reincidência, auxiliando o juiz na definição da pena a ser aplicada ao acusado: quanto maior o risco estimado de que o indivíduo volte a delinquir, maior tende a ser a pena recomendada (Zanon; Branco; Silva, 2025, p 190).

Conforme adverte Zuboff, "algoritmos treinados com dados históricos tendem a replicar e amplificar os padrões de desigualdade presentes na sociedade, criando um ciclo de retroalimentação de injustiças sob a aparência de neutralidade técnica" (Zuboff, 2021, p. 298).

Outro ponto de tensão ética refere-se à atribuição de responsabilidade moral e jurídica. Em um cenário em que decisões judiciais são influenciadas por algoritmos, quem deve ser responsabilizado por erros, injustiças ou violações de direitos fundamentais? A inteligência artificial funcionaria como instrumento auxiliar ao intérprete, permitindo-lhe fundamentar suas decisões, ao aplicar o Direito, em parâmetros normativos respaldados por evidências. Ademais, sob uma ótica normativa, o intérprete não poderia simplesmente desconsiderar de modo irresponsável as evidências apresentadas pela IA, sem assumir o ônus argumentativo de sustentar o padrão normativo previamente estabelecido, em consonância com a lógica do *stare decisis et non quieta non movere* (Morais, 2020, p. 57-58).

Dessa sorte, a responsabilidade constitui o núcleo ético da atividade jurisdicional; se não for possível identificar quem responde por um ato decisório, o Direito perde sua base moral e deixa de cumprir sua função de garantia. A multiplicidade de agentes envolvidos — desenvolvedores, magistrados, instituições públicas e fornecedores de tecnologia — gera uma diluição da responsabilidade, dificultando a imputação de culpa ou dolo, o que compromete o princípio da *accountability* e o próprio ideal de justiça. Zuboff ressalta que "estabelecer limites claros sobre o que pode ser delegado a máquinas é essencial para evitar a diluição gradual de responsabilidades e a erosão da legitimidade institucional" (Zuboff, 2021, p. 415).

Somado a isso, há o risco da desumanização da decisão judicial. Segundo Bichara

e Brito, o recurso a dados históricos de prisões e condenações tende a reproduzir as discriminações neles já presentes, especialmente quando utilizado para decisões sobre fiança e definição de sentenças. A incorporação de sistemas de IA nessas etapas pode conduzir à desumanização da atividade jurisdicional, uma vez que afasta uma análise integral da pessoa envolvida, fragilizando a confiança social no processo. Ademais, a falta de transparência dos algoritmos e a inexistência de mecanismos precisos de responsabilização podem intensificar desigualdades já existentes (Bichara; Brito, 2024, pp. 12-13).

A dependência crescente de sistemas automatizados tende a reduzir o papel crítico e interpretativo do julgador, conduzindo ao fenômeno *automation bias*, em que decisões são tomadas de forma acrítica com base nas sugestões da máquina. "O viés de automação refere-se à tendência de, em certas situações, os seres humanos confiarem em sistemas de apoio à decisão baseados em IA para além do que seria adequado, considerando o nível real de confiabilidade dos algoritmos" (Horowitz; Kahn, 2024, p. 3, tradução livre).[4] Dworkin, ao analisar a interpretação construtiva que deve ser realizada pelo juiz, demonstra que julgar constitui um ato de interpretação com dimensão ética, e não uma operação automática de enquadramento lógico: "Em linhas gerais, a interpretação construtiva é uma questão de impor um propósito a um objeto ou prática, a fim de torná-lo o melhor exemplo possível da forma ou gênero aos quais se imagina que pertençam. [...] Do ponto de vista construtivo, a interpretação criativa é um caso de interação entre propósito e objeto." (Dworkin, 2021, pp. 63-64).

Ao abdicar de sua autonomia ética, o juiz converte-se em mero executor de resultados computacionais, o que enfraquece a legitimidade da jurisdição e ameaça o próprio conceito de justiça como virtude moral. Por fim, o debate ético sobre a IA na decisão jurídica deve considerar a preservação da dignidade da pessoa humana. O uso de ferramentas de predição ou classificação automatizada no Judiciário pode reduzir o indivíduo a uma abstração estatística, ignorando sua singularidade existencial. Teigão e Fogaça ressaltam que "apesar das potencialidades e modernização, o uso de sistemas algorítmicos deve submeter-se a controles institucionais rigorosos, com supervisão humana e mecanismos de auditabilidade, para preservar a confiança e legitimidade das decisões judiciais" (Teigão; Fogaça, 2025, p. 138).

Portanto, a ética da inteligência artificial deve assegurar a centralidade da pessoa humana, de modo que a tecnologia sirva ao homem e não o contrário. A dignidade deve

[4] Original: "Automation bias refers to the tendency of humans, in some situations, to rely on AI decision aids above and beyond the extent to which they should, given the reliability of the algorithms".

funcionar como limite moral e normativo da automação decisória, reafirmando o princípio de que a justiça não é apenas uma operação racional, mas também um exercício de empatia e prudência. Entre a decisão jurídica dada pelo juízo e aquela franqueada total ou parcialmente à IA, Morais analisa que o "robô" não pode ser juiz: "[...] somente o ser humano pode aplicar o Direito. Esse e o argumento do antropocentrismo jurídico. Entretanto, objeta-se a relativização desse argumento, admitindo-se que a IA pode aplicar o Direito em determinadas situações independentemente da interveniência humana" (Morais, 2020, p. 58). Isso exige o desenvolvimento de uma ética da decisão automatizada, capaz de integrar valores humanos ao desenho e à utilização dos sistemas de IA no Direito. Tal ética deve envolver transparência algorítmica, auditoria independente, mecanismos de revisão humana e políticas de governança que garantam a responsabilidade e a proteção dos direitos fundamentais. A IA pode, assim, contribuir para uma justiça mais eficiente, sem renunciar àquilo que a torna verdadeiramente justa: o respeito pela dignidade e pela moralidade do ser humano.

3 PERSPECTIVAS REGULATÓRIAS E CAMINHOS PARA A ÉTICA DA DECISÃO AUTOMATIZADA

O enfrentamento dos desafios éticos decorrentes da aplicação da Inteligência Artificial (IA) no processo decisório jurídico exige a construção de um arcabouço normativo fortalecido e interdisciplinar, capaz de compatibilizar inovação tecnológica e proteção dos direitos fundamentais. O Direito, enquanto instrumento de limitação do poder e de promoção da justiça, deve assumir papel protagonista na regulação do uso da IA, não apenas como reação a seus efeitos, mas como instância preventiva e orientadora de seu desenvolvimento ético e social.

Em nível internacional, destacam-se as iniciativas da União Europeia, que vem liderando o movimento global de regulamentação ética da IA. O *AI Act*, aprovado em 2024 pelo Parlamento Europeu, estabelece um modelo regulatório baseado em riscos graduais, classificando sistemas de IA conforme o impacto potencial sobre os direitos humanos e as liberdades civis (European Commission, 2024). Os sistemas aplicados no campo judicial são enquadrados como de alto risco, exigindo requisitos rigorosos de transparência, rastreabilidade e supervisão humana significativa (*meaningful human oversight*). Vasconcelos, Jörke, Grunde-Mclaughlin, Gerstenberg e Krishna (2023) demonstram em pesquisa publicada que explicações claras e bem estruturadas podem mitigar o viés de automação, ao incentivar uma participação mais reflexiva e uma supervisão humana efetiva sobre as decisões da IA. Além disso, os autores observam que a explicabilidade é

especialmente valorizada em situações que envolvem tarefas complexas, quando as justificativas apresentadas são diretas e existem incentivos, como benefícios financeiros.

Essa abordagem evidencia a preocupação em impedir a automatização cega de decisões que afetam a vida das pessoas, reafirmando a necessidade de preservar o julgamento ético como dimensão essencial da justiça. Sugere-se que explicações bem elaboradas podem reduzir o viés de automação ao promover um engajamento mais crítico e uma supervisão humana mais significativa sobre as recomendações da IA (Vasconcelos; Jörke; Grunde-Mclaughlin; Gerstenberg; Krishna, 2023).

No Brasil, embora ainda incipiente, há avanços normativos importantes. O Conselho Nacional de Justiça (CNJ) editou a Resolução n. 615/2024 (que revogou a Resolução n. 332/2020), que dispõe sobre a ética, a governança e a transparência no uso da inteligência artificial no Poder Judiciário. Essa resolução estabelece princípios como explicabilidade, responsabilidade, equidade, segurança e respeito aos direitos fundamentais, determinando que a IA deve atuar como instrumento de apoio à decisão, e não como substituta da função jurisdicional humana. Tal diretriz reflete uma compreensão ética avançada: a tecnologia pode aprimorar a eficiência e a gestão judicial, mas não deve assumir o papel moral do julgador (Conselho Nacional de Justiça, 2024).

Merece destaque o princípio da explicabilidade. Nesse sentido, Macedo, Dias e Cavalcante explicam que, com a elevação dos padrões aplicáveis aos sistemas de IA, tornou-se necessária uma abordagem que valorizasse a explicabilidade dos modelos, dando origem ao conceito de Inteligência Artificial Explicável (xAI). Tal conceito diz respeito a um conjunto de métodos destinados a tornar a IA mais transparente e compreensível ao público, permitindo que os usuários compreendam os fundamentos, os motivos e o processo pelo qual um sistema chegou a determinada conclusão, decisão ou previsão. Além disso, a explicabilidade contribui para a identificação e mitigação de vieses, favorecendo o desenvolvimento de sistemas mais responsáveis e com funcionamento mais controlável (Macedo; Dias; Cavalcante, 2025, p. 54)

Entretanto, a efetividade desses marcos regulatórios depende da implementação de mecanismos concretos de controle e auditoria algorítmica. Percebe-se, assim, que A regulação do uso de inteligência artificial no âmbito do Judiciário precisa ser articulada com mecanismos de auditoria pública e transparência, assegurando a efetiva prestação de contas e, consequentemente, a legitimidade democrática das decisões automatizadas. A criação de comitês de ética tecnológica e de protocolos de transparência algorítmica é medida necessária para assegurar o cumprimento desses princípios utilizam (Bichara; Brito, 2024, p.14).

Além disso, deve-se promover a alfabetização digital de magistrados e servidores,

de forma que compreendam o funcionamento, os limites e os riscos éticos das ferramentas que utilizam (Bichara; Brito, 2024, p.14). No plano filosófico-jurídico, a regulação da IA deve ser orientada pelo princípio da tecnologia a serviço da humanidade (Müller, 2025, p. 10). Isso significa reconhecer que o valor ético da tecnologia não reside em sua capacidade de substituir o humano, mas de ampliar sua autonomia moral e racional. Para Floridi:

[...] to use a metaphor already introduced above, the best way to catch the technology train is not to chase it, but to be at the next station. In other words, we need to anticipate and steer the ethical development of technological innovation. And we can do this by looking at what is actually feasible, privileging, within this, what is environmentally sustainable, then what is socially acceptable and then, ideally, choose what is socially preferable (Floridi, 2018, p. 7).[5]

Não se trata apenas de uma questão técnica: a regulação da IA é um imperativo ético, que define as condições sob as quais confiamos decisões relevantes a sistemas artificiais. Após analisar a questão ética da IA na Justiça no Direito Comparado, Teigão e Fogaça destacam que "em todas essas jurisdições, o uso da IA é condicionado à existência de estruturas claras de governança, à supervisão permanente de agentes humanos e à realização de avaliações de impacto regulatório e ético" (Teigão; Fogaça, 2025, p. 153). A regulação deve, portanto, preservar a supervisão humana significativa, conceito essencial à ética da automação. Essa supervisão não se reduz ao ato formal de validação, mas implica participação crítica, interpretativa e responsável do juiz nas etapas do processo decisório.

Outro ponto relevante é a necessidade de garantir a explicabilidade e a auditabilidade das decisões baseadas em IA. Nessa linha, direitos à revisão e à explicação fundamentam-se, sobretudo, na necessidade de incorporar um elemento humano ao processo decisório automatizado, especialmente no caso de sistemas de IA. Tal inserção visa possibilitar que a decisão seja submetida a um controle de legitimidade baseado em critérios compreensíveis e passíveis de contestação por pessoas. Nesse sentido, tais direitos funcionam como instrumentos destinados a enfrentar o problema da confiança na máquina. Assim, compreende-se que outros parâmetros de legitimação também desempenham papel relevante, podendo relativizar tanto a forma quanto o grau de participação humana exigida nos processos decisórios automatizados (Doneda; Wimmer, 2021, pp. 399-400).

[5] Tradução livre: [...] para usar uma metáfora já mencionada anteriormente, a melhor forma de "pegar o trem da tecnologia" não é correr atrás dele, mas estar na próxima estação. Em outras palavras, precisamos antecipar e orientar o desenvolvimento ético da inovação tecnológica. E podemos fazer isso observando o que é efetivamente viável, privilegiando, dentre essas possibilidades, aquilo que é ambientalmente sustentável, depois o que é socialmente aceitável e, idealmente, escolher o que é socialmente preferível

Nesse contexto, infere-se que a ética na inteligência artificial depende da explicabilidade, pois é ela que possibilita o exame moral e jurídico das decisões automatizadas. No âmbito judicial, isso significa que cada decisão assistida por IA deve ser rastreável, com registro claro dos parâmetros, fontes de dados e inferências utilizadas pelo sistema. Tais medidas concretizam o dever de motivação e asseguram o controle recursal, pilares do devido processo legal. Sob a perspectiva institucional, recomenda-se o desenvolvimento de um Código de Ética para o uso de IA no sistema de justiça, inspirado em experiências internacionais como a *Ethical Charter on the Use of Artificial Intelligence in Judicial Systems*, elaborada pelo Conselho da Europa em 2018 (Council of Europe, 2018, pp. 7-12). Esse documento reconhece cinco princípios fundamentais: respeito aos direitos humanos, não discriminação, qualidade e segurança, transparência e controle pelo usuário.

No contexto brasileiro, tal código pode consolidar diretrizes de conduta para os profissionais do direito, contribuindo para um ecossistema de governança ética e responsável. Entretanto, mecanismos normativos que podem servir mecanismos de preservação da conduta ética dos envolvidos no campo de decisões jurídicas. Trata-se do Marco Civil da Internet (Lei 12.965/14), que estabeleceu diretrizes importantes para o uso da *internet*, incluindo a proteção de dados pessoais e a privacidade dos usuários, e a Lei Geral de Proteção de Dados (Lei 13.709/18), que reforçou esses princípios ao regular o tratamento de dados pessoais, garantindo transparência e segurança na coleta e uso dessas informações (Brasil, 2025).

Cabe destacar os desafios relacionados à hipernormatização jurídica com a utilização de Inteligência artificial no Direito, exigindo cuidado na programação e na aplicação automatizada das decisões jurídicas. Nesse sentido, Morais destaca que:

A hipernormatização jurídica pode ser um problema do uso da AI no Direito, exigindo do jurista uma atenção sobre a forma pela qual a programação computacional é realizada com base em decisões jurídicas sumuladas. Essa atenção envolve argumentar sempre privilegiando os fatores jurídicos e fáticos determinantes para a formalização de premissas que serão posteriormente utilizadas para a aplicação de forma automatizada do Direito pela IA. Quando se fala na aplicação do Direito, há sempre a preocupação de que este ato seja realizado pelo ser humano. Esse é o argumento do antropocentrismo jurídico. Entende-se que tal argumento não é absoluto, sendo possível que em determinados casos conhecidos como repetitivos se reconheça a possibilidade de aplicação do Direito pela IA. Por fim, a superação da crise ética da decisão jurídica automatizada depende não apenas de normas positivas, mas de uma mudança de paradigma. O Direito deve abandonar a postura reativa diante da tecnologia e adotar uma ética da coevolução, na qual o

desenvolvimento tecnológico e o progresso moral caminhem em conjunto (Morais, 2020, p. 58).

A inteligência artificial deve ser incorporada ao Judiciário como extensão da racionalidade humana, e não como substituta de sua consciência ética (Teigão; Fogaça, 2025, p. 146). Essa visão propõe uma convivência harmônica entre razão técnica e razão moral, reafirmando o papel do juiz como guardião da justiça em tempos de algoritmos. Outrossim, a ética digital impacta o desenvolvimento tecnológico ao orientar a sociedade sobre o que é considerado moralmente correto e socialmente aceitável. Conforme dispõe Floridi:

In particular, digital ethics, with its values, principles, choices, recommendations and constrains already influences the world of technology much more than any other force. This is because the evaluation of what is morally good, right or necessary shapes public opinion—hence the socially (Floridi, 2018, p. 7).[6]

Assim, a consolidação de um marco ético-regulatório para a IA no Direito requer o equilíbrio entre inovação, responsabilidade e humanidade. Conforme expõe Barreto Junior:

A inteligência artificial, ao atuar como agente decisório em situações que afetam diretamente a vida e os direitos fundamentais dos cidadãos, demanda a imposição de limites éticos claros e precisos. A opacidade inerente aos algoritmos e a capacidade das máquinas de processar volumes incomensuráveis de dados pessoais devem ser contrabalançadas por um sistema jurídico que assegure a transparência e a *accountability* no uso dessas tecnologias. É inaceitável que decisões automatizadas, muitas vezes invisíveis ao olho humano, perpetuem desigualdades e consolidem discriminações estruturais (Barreto Junior, 2024, p. 93).

O desafio do século XXI consiste em garantir que a eficiência tecnológica não se converta em indiferença moral, e que o futuro da decisão jurídica permaneça fiel ao princípio fundamental que sustenta o próprio Estado de Direito: a dignidade da pessoa humana.

CONCLUSÃO

A análise desenvolvida ao longo deste estudo evidencia que a introdução da

[6] Tradução livre: Em particular, a ética digital, com seus valores, princípios, escolhas, recomendações e limitações, já influencia o mundo da tecnologia muito mais do que qualquer outra força. Isso ocorre porque a avaliação do que é moralmente bom, correto ou necessário molda a opinião pública — e, portanto, o que é socialmente aceitável.

Inteligência Artificial (IA) no processo decisório jurídico não se limita a uma questão de eficiência ou inovação tecnológica, mas coloca em crise a própria essência ética do ato de julgar.

A pesquisa partiu da hipótese de que a automação crescente das decisões judiciais tensiona os fundamentos morais, axiológicos e de responsabilidade que sustentam o Direito, transformando o julgamento em um espaço suscetível à opacidade, aos vieses e à diluição da responsabilidade humana. Como se viu, o estudo confirma que tais riscos não são hipotéticos: a IA, quando utilizada sem parâmetros éticos claros, pode comprometer a transparência, a imparcialidade e a dignidade da pessoa humana, valores estruturantes do Estado Democrático de Direito.

O estudo destacou que os desafios éticos surgem em múltiplas dimensões. Primeiramente, a opacidade algorítmica dificulta a compreensão e a contestação das decisões assistidas por IA, fragilizando o dever de motivação e a *accountability*. Em segundo lugar, a reprodução de vieses e desigualdades históricas evidencia que a neutralidade tecnológica é ilusória e que o uso indiscriminado de algoritmos pode consolidar injustiças sociais. Por fim, a diminuição da autonomia decisória do julgador e a potencial desumanização da jurisdição indicam que a mera eficiência não pode substituir a reflexão moral, o discernimento ético e a responsabilidade individual do magistrado.

Diante dessas problemáticas, a pesquisa apontou caminhos para a superação da crise ética, destacando a importância de um marco regulatório eficiente e interdisciplinar, que combine legislação, resoluções administrativas, auditoria de sistemas e códigos de ética específicos. Iniciativas internacionais, como a Diretiva Europeia de IA (2024), e nacionais, como a Resolução CNJ n. 615/2024, fornecem referências concretas para a construção de políticas que assegurem a transparência, a explicabilidade, a supervisão humana significativa e o respeito aos direitos fundamentais.

Além disso, a ética da decisão automatizada deve ser entendida como um conjunto de práticas contínuas que orientam o desenvolvimento, a implementação e a utilização da IA, garantindo que a tecnologia permaneça subordinada à racionalidade moral humana. Em síntese, a crise da decisão jurídica ética não se limita a um problema técnico ou operacional, mas é uma questão epistemológica e moral, que exige a integração entre Direito, Filosofia e Ética. A superação desse desafio depende de reafirmar o papel do juiz como agente moral e responsável, de construir sistemas de IA transparentes e auditáveis, e de instituir normas e práticas que garantam que a inovação tecnológica sirva à justiça, e não à eficiência desprovida de valores.

A pesquisa reforça, assim, a necessidade de pensar a Inteligência Artificial como instrumento de ampliação da racionalidade ética e não como substituta do julgamento

humano, preservando, em última instância, a dignidade da pessoa e os princípios fundamentais que estruturam o Estado de Direito.

REFERÊNCIAS

ANGWIN, Julia; LARSON, Jeff; MATTU, Surya; KIRCHNER, Lauren. Machine Bias. *In*: **ProPublica**, 2016. Disponível em: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. Acesso em: 28 out. 2025.

ARISTÓTELES. **Ética a Nicômaco**. Tradução e notas: Edson Bini. São Paulo: Edipro, 2018.

BARRETO JUNIOR, Irineu Francisco. Ética e Inteligência Artificial: desafios na modulação e regulação dos algoritmos. *In*: **Cadernos Adenauer**, ano XXV (2024), n. 3, Inovações Tecnológicas e Democracia. Rio de Janeiro: Fundação Konrad Adenauer, dezembro 2024, pp. 81-95.

BICHARA, Anderson de Andrade; BRITO, Fabiana Amaro de. Desafios éticos ao uso da inteligência artificial no sistema de justiça criminal. *In*: **Boletim IBCCRIM**, v. 32, n. 382, pp. 11-14, 2024.

BRASIL. **Constituição da República Federativa do Brasil, de 05 de outubro de 1988**. Brasília, DF, 2025.

BRASIL. **Lei n. 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados (LGPD). Brasília, DF. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 28 out. 2025.

BRASIL. **Marco Civil da Internet.** Lei n. 12.965, de 23 de abril de 2024. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm. Acesso em: 28 out. 2025.

CONSELHO NACIONAL DE JUSTIÇA. **Resolução n. 615, de 11 de março de 2025**. Dispõe sobre a ética, a transparência e a governança na produção e no uso de Inteligência Artificial no Poder Judiciário e dá outras providências. Disponível em: https://atos.cnj.jus.br/atos/detalhar/3429. Acesso em: 28 out. 2025.

COUNCIL OF EUROPE. **European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their Environment**. Strasbourg, 2018. Disponível em: https://www.europarl.europa.eu/cmsdata/196205/COUNCIL%20OF%20EUROPE%20-%20European%20Ethical%20Charter%20on%20the%20use%20of%20AI%20in%20judicial%20systems.pdf. Acesso em: 28 out. 2025.

DONEDA, Danilo; WIMMER, Miriam. "Falhas de IA" e a Intervenção Humana em Decisões Automatizadas: Parâmetros para a Legitimação pela Humanização. **Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa (IDP)**. Dossiê – Inteligência Artificial, Ética e Epistemologia. Brasília, v. 18, n. 100, pp. 374-406, out./dez. 2021.

DWORKIN, Ronald. **O Império do Direito**. Tradução Jefferson Luiz Camargo. São Paulo: Martins Fontes - selo Martins, 2014, 3ª edição - reimpressão 2021.

EUROPEAN COMMISSION. **Artificial Intelligence Act (AIA)**. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)**.** Brussels, July 12, 2024. Disponível em: https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=OJ:L_202401689. Acesso em: 28 out. 2025.

FLORIDI, Luciano. **Soft Ethics and the Governance of the Digital**. *In*: Philosophy & Technology, EUA, pp. 1-7, February 2018.

FLORIDI, Luciano; COWLS, Josh. A Unified Framework of Five Principles for AI in Society. Harvard **Data Science Review (HDSR), EUA,** v. 3, n. 1, Jul 01, 2019.

HABERMAS, Jürgen. **Facts and Norms: Contributions to a Discourse Theory of Law and Democracy.** Translated by William Rehg. Second printing. Cambridge, Massachusets: The MIT Press, 1996.

HOROWITZ, Michael C.; KANH, Lauren Kahn. Bending the Automation Bias Curve: A Study of Human and AI-Based Decision Making in National Security Contexts. International Studies Quarterly. **International Studies Quarterly**, volume 68, issue 2, June 2024, sqae020, EUA, Disponível em: https://doi.org/10.1093/isq/sqae020. Acesso em: 28 out. 2025.

MACEDO, Gabriela de Araújo; DIAS, Lucas Cerqueira Machado; CAVALCANTE, Laura Galvão Marques. Ética e Inteligência Artificial: Os Desafios do Princípio da Explicabilidade à Luz do PL 2.338/23. *In*: **Inteligência Artificial aplicada:** soluções para um mundo automatizado [livro digital]. Org.: Ernane Rosa Martins. Guarujá-SP: Científica Digital, pp. 45-61, 2025.

MORAIS, Fausto Santos de. Três Questões sobre a Integração da Inteligência Artificial no Direito. *In*: **Direito, Democracia e Sustentabilidade**. Anuário do Programa de Pós-Graduação *Stricto Sensu* em Direito da Faculdade Meridional. Orgs.: Márcio Ricardo Staffen e Neuro José Zambam. Cruz Alta: Ilustração, pp. 50-59, 2020.

MÜLLER, Vicent C. Ethics of Artificial Intelligence and Robotics. **Stanford Encyclopedia of Philosophy**, EUA, pp. 1-23, 2025.

PERELMAN, Chaim. **Ética e direito**. Tradução: Maria Ermantina Galvão G. Pereira. São Paulo: Martins Fontes, 1996.

SIQUEIRA, Dirceu Pereira; MORAIS, Fausto Santos de Morais; SANTOS, Marcel Ferreira dos. Inteligência artificial e jurisdição: dever analítico de fundamentação e os limites da substituição dos humanos por algoritmos no campo da tomada de decisão judicial. **Sequência: estudos jurídicos e políticos**, Florianópolis, vol. 43, n. 91, pp. 1-36, 2022.

SOUSA, Susana Aires de. "Não fui eu, foi a máquina": Teoria do crime, responsabilidade e inteligência artificial. *In*: **A Inteligência Artificial no Direito Penal**. Coord.: Anabela Miranda Rodrigues. Coimbra: Almedina, pp. 59-93, 2020.

TEIGÃO, Rafael Coninck; FOGAÇA, Leonardo de Andrade Ferraz. O uso ético e responsável da inteligência artificial no Judiciário brasileiro: impactos para o

jurisdicionado e balizas regulatórias. **Revista Jurídica Gralha Azul – TJPR**, v. 1 n. 28, pp. 136-162. mar/2025 a abril/2025.

VALLE, Vivian Lima López; FUENTES i GASÓ, Josep Ramón; AJUS, Attílio Martins. Decisão judicial assistida por inteligência artificial e o Sistema Victor do Supremo Tribunal Federal. **Revista de Investigações Constitucionais**, Curitiba, vol. 10, n. 2, e252, maio/ago., pp. 1-38, 2023.

VASCONCELOS, Helena; JÖRKE, Matthew; GRUNDE-MCLAUGHLIN, Madeleine; GERSTENBERG, Tobias; BERNSTEIN, Michael S.; KRISHNA, Ranjay. Explanations can reduce overreliance on ai systems during decision-making. **Proceedings of the ACM on Human-Computer Interaction**, 7(CSCW1), pp. 1-38, 2023.

ZANON, Orlando; BRANCO, Matheus de Andrade; SILVA, Pollyana Maria da. Inteligência artificial e vieses morais na tomada de decisão. **Revista da AGU**, Brasília/DF, v. 24, n. 1, pp. 176-197, 2025.

ZUBOFF, Shoshana. **A era do capitalismo de vigilância**: A luta por um futuro humano na nova fronteira do poder. Tradução: George Schlesinger. Rio de Janeiro: Intrínseca, 2021.