



**UNIVERSIDADE FEDERAL DO CEARÁ**  
**FACULDADE DE DIREITO**

**JOÃO VICTOR NASCIMENTO LIMA**

**ÉTICA DA RESPONSABILIDADE E INTELIGÊNCIA ARTIFICIAL: UM OLHAR A  
PARTIR DE HANS JONAS**

**FORTALEZA**

**2025**

**JOÃO VICTOR NASCIMENTO LIMA**

ÉTICA DA RESPONSABILIDADE E INTELIGÊNCIA ARTIFICIAL: UM OLHAR A  
PARTIR DE HANS JONAS

Monografia apresentada ao curso de  
Direito da Universidade Federal do Ceará,  
como requisito parcial para obtenção do  
título de bacharel em Direito.

Orientador: Dra. Prof.<sup>a</sup> Fernanda Claudia  
Araújo da Silva.

Fortaleza

2025

---

L1f LIMA, João Victor Nascimento.

ÉTICA DA RESPONSABILIDADE E INTELIGÊNCIA ARTIFICIAL: UM OLHAR A PARTIR DE HANS JONAS / João Victor Nascimento LIMA. – 2025.

52 f. : il. color.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Faculdade de Direito, Curso de Direito, Fortaleza, 2025.

Orientação: Dra. Prof.<sup>a</sup> Fernanda Claudia Araújo da Silva. 1. inteligência artificial. 2. Ética da responsabilidade 3. Regulamentação. I. Título. CDD 340

---

JOÃO VICTOR NASCIMENTO LIMA

ÉTICA DA RESPONSABILIDADE E INTELIGÊNCIA ARTIFICIAL: UM OLHAR A  
PARTIR DE HANS JONAS

Monografia apresentada ao curso de  
Direito da Universidade Federal do Ceará,  
como requisito parcial para obtenção do  
título de bacharel em Direito.

Aprovada em: \_\_\_/\_\_\_/\_\_\_.

BANCA EXAMINADORA

---

Dra. Prof.<sup>a</sup> Fernanda Claudia Araújo da Silva (Orientadora)  
Universidade Federal do Ceará (UFC)

---

Prof.<sup>o</sup> Diego Gomes da Silva  
Universidade Federal do Ceará (UFC)

---

Prof.<sup>o</sup> José Gutemberg de Sousa Rodrigues Júnior  
Universidade Federal da Paraíba (UFPB)

Este trabalho é dedicado à Deus e à minha mãe, pois, graças a eles que hoje posso concluir este glorioso Curso de Direito.

## **AGRADECIMENTOS**

A Deus, fonte inesgotável de luz e sabedoria, que me concedeu forças nos momentos de incerteza e perseverança diante dos desafios ao longo desta jornada acadêmica.

Aos meus professores, verdadeiros pilares nesta trajetória, cuja dedicação e conhecimento não apenas enriqueceram minha formação acadêmica, mas também despertaram em mim o desejo contínuo pelo aprendizado e pelo aprimoramento profissional.

Aos meus familiares, que sempre acreditaram no meu potencial e me apoiaram incondicionalmente, compreendendo minhas ausências e me incentivando a seguir em frente, mesmo nos momentos mais árduos.

Aos amigos, que compartilharam comigo essa jornada, seja nos momentos de estudo, de trocas de conhecimento ou nos instantes de descontração necessários.

Por fim, a todos que, de alguma forma, contribuíram para a realização deste trabalho, seja com palavras de incentivo, auxílio técnico ou inspiração, deixo minha sincera gratidão. Este TCC representa não apenas uma conquista pessoal, mas também o reflexo do apoio e das contribuições daqueles que caminharam comigo nesta trajetória.

"O aspecto mais triste da vida atual é que a ciência ganha em conhecimento mais rapidamente que a sociedade em sabedoria."

Isaac Asimov

## RESUMO

Objetiva-se explorar os desafios éticos e regulatórios que surgem com a aplicação da Inteligência Artificial no Direito, utilizando a Filosofia do Direito proposta por Hans Jonas como base normativa, sobretudo considerando que tal sistema de Inteligência Artificial está sendo cada vez mais implementada em diversas áreas do campo jurídico. Levanta-se também questões éticas importantes, como a opacidade dos algoritmos, a reprodução de vieses algorítmicos e a responsabilidade legal em decisões automatizadas, discutindo-se como todos estes problemas podem ser influenciados pela perspectiva oferecida por Hans Jonas. A pesquisa baseia-se em uma revisão bibliográfica abrangente e na análise de estudos de caso para discutir como a Ética da Responsabilidade proposta por Hans Jonas pode servir de baluarte para regular o uso de Inteligência Artificial no sistema jurídico de forma justa e ética. O trabalho propõe um quadro regulatório que enfatiza a transparência, a responsabilidade e a equidade, concluindo que a perspectiva sintetizada por Hans Jonas pode servir de substrato norteador da necessária regulamentação da Inteligência Artificial.

**Palavras-chave:** Inteligência artificial; Ética da Responsabilidade; Regulamentação.

## **ABSTRACT**

The objective is to explore the ethical and regulatory challenges that arise with the application of Artificial Intelligence in Law, using the Philosophy of Law proposed by Hans Jonas as a normative foundation, especially considering that such Artificial Intelligence systems are increasingly being implemented in various areas of the legal field. Important ethical issues are also raised, such as algorithmic opacity, the reproduction of algorithmic biases, and legal responsibility in automated decisions, discussing how all these problems can be influenced by the perspective offered by Hans Jonas. The research is based on a comprehensive literature review and case study analysis to examine how the ethics of responsibility proposed by Hans Jonas can serve as a cornerstone for regulating the use of Artificial Intelligence in the legal system in a fair and ethical manner. The study proposes a regulatory framework that emphasizes transparency, accountability, and fairness, concluding that the perspective synthesized by Hans Jonas can serve as a guiding foundation for the necessary regulation of Artificial Intelligence.

**Keywords:** Artificial Intelligence; Philosophy of Law; Ethics; Regulation.

## SUMÁRIO

<b>1. Introdução.....</b>	<b>11</b>
<b>2. Fundamentos Teóricos da Inteligência Artificial e sua Aplicação no Direito.....</b>	<b>13</b>
<b>2.1. Conceito de Inteligência Artificial e Aprendizado de Máquina.....</b>	<b>13</b>
<b>2.2. Aplicações da IA no Campo Jurídico.....</b>	<b>15</b>
<b>2.3. Vantagens da Utilização da IA no Direito.....</b>	<b>17</b>
<b>2.4. Limitações e Problemas Inerentes à Implementação da IA no Direito.....</b>	<b>19</b>
<b>3. Desafios Éticos da Inteligência Artificial no Sistema Jurídico.....</b>	<b>23</b>
<b>3.1. Imparcialidade e Discriminação Algorítmica.....</b>	<b>23</b>
<b>3.2. Transparência e o Problema da "Caixa-Preta".....</b>	<b>24</b>
<b>3.3. Responsabilidade Legal e Ética no Uso da IA.....</b>	<b>27</b>
<b>3.4. Privacidade e Proteção de Dados.....</b>	<b>30</b>
<b>3.5. Autonomia e Controle Humano nas Decisões Automatizadas.....</b>	<b>32</b>
<b>4. A Perspectiva De Hans Jonas Como Base Dogmática Para A Regulação da IA No Direito.....</b>	<b>36</b>
<b>4.1. A Ética da Responsabilidade de Hans Jonas.....</b>	<b>36</b>
<b>4.1.1 A necessidade da modernização da definição de ética.....</b>	<b>37</b>
<b>4.1.2 Fundamentos da Ética da Responsabilidade.....</b>	<b>38</b>
<b>4.2. A aplicação da Ética da Responsabilidade na regulação da aplicação da inteligência artificial no Direito.....</b>	<b>40</b>
<b>5. Considerações.....</b>	<b>43</b>
<b>Referências.....</b>	<b>45</b>

## 1. INTRODUÇÃO

A Inteligência Artificial (IA) tem se consolidado como uma inovação transformadora em diversas áreas, inclusive, seu impacto no campo do Direito é particularmente significativo. Em tal área, a IA está sendo aplicada em atividades que variam desde a análise de contratos e documentos jurídicos até a tomada de decisões judiciais assistida por algoritmos, sob a promessa de garantir de maior eficiência, celeridade processual e redução de custos, além da possibilidade de reduzir vieses humanos.

Inobstante, tal cenário faz emergir o questionamento quanto aos riscos e desafios éticos que precisam ser cuidadosamente considerados e abordados para a utilização da IA no Direito, tornando relevante a ponderação quanto à regulamentação sobre o assunto.

Nesse contexto, a Filosofia do Direito, em especial a perspectiva da Ética da Responsabilidade proposta por Hans Jonas, surge como substrato essencial para instruir uma base normativa ética e eficaz da regulamentação da IA no Direito, uma vez que nela podemos encontrar princípios e valores capazes de orientar a criação de marcos regulatórios próximos ao ideal de justiça.

Dessa forma, o argumento principal deste trabalho é analisar se a perspectiva proposta por Hans Jonas fornece uma base dogmática adequada para a regulamentação ética e eficaz da Inteligência Artificial aplicada ao Direito, de modo que, através dessa lente filosófica, seja possível desenvolver um arcabouço regulatório que promova o uso da IA de maneira que respeite as garantias individuais dos cidadãos, ao passo que em que garante a segurança quanto aos algoritmos utilizados.

Para explorar essa perspectiva, este trabalho pretende abordar o tema através de uma revisão de literatura abrangente, que incluirá a análise artigos e publicações relacionados aos ensinamentos de Hans Jonas, outrossim o estudo de regulações emergentes e opiniões de especialistas no campo do Direito Digital e da Inteligência Artificial. Ademais, o estudo examinará estudos recentes que exploram a integração da IA em sistemas jurídicos, além dos desafios éticos resultantes dessa inovação.

O trabalho será dividido em três capítulos principais. O primeiro capítulo apresentará os fundamentos teóricos da Inteligência Artificial e sua

aplicação no Direito, discutindo conceitos básicos e implicações práticas.

O segundo capítulo, por sua vez, analisará os desafios éticos que a IA impõe ao sistema jurídico, como a discriminação algorítmica, a falta de transparência e a responsabilidade jurídica, refletindo como essas problemáticas refletem nas garantias dos cidadãos.

O terceiro e último capítulo abordará como a perspectiva filosófica da Ética da Responsabilidade de Hans Jonas pode fornecer diretrizes concretas para uma regulamentação ética da Inteligência Artificial no Direito.

O objetivo deste trabalho é fornecer uma visão abrangente de como a perspectiva de Hans Jonas pode fornecer uma base normativa essencial para uma regulamentação ética da Inteligência Artificial no campo jurídico.

As questões de pesquisa que guiarão esta análise são: Como a perspectiva de Hans Jonas pode contribuir para o desenvolvimento de um marco regulatório ético para o uso da IA no Direito? Quais são os desafios éticos específicos que surgem da aplicação da IA no sistema jurídico, e como eles podem ser abordados por meio de uma abordagem filosófica de Hans Jonas? De que maneira a filosofia de Hans Jonas pode garantir a transparência, a responsabilidade e a justiça na utilização de IA, evitando práticas desabonem os Direitos Fundamentais dos Cidadãos e contrárias a um sistema legal mais equitativo e eficaz?

## **2. FUNDAMENTOS TEÓRICOS DA INTELIGÊNCIA ARTIFICIAL E SUA APLICAÇÃO NO DIREITO**

A IA, conforme a definição do criador do termo, John McCarthy (1998), é a “a ciência e engenharia de produzir sistemas, inteligentes”. Ou seja, tal ferramenta consiste em uma área da ciência cujo propósito é estudar, desenvolver e utilizar máquinas e sistemas para realizarem atividades humanas de modo autônoma, de modo a imitar a capacidade cognitiva humana por meio de algoritmos (Barbosa, 2023).

Seu desenvolvimento, impulsionado pelo avanço da capacidade computacional e pelo aprimoramento dos algoritmos, tem possibilitado a automação de tarefas antes exclusivamente humanas, promovendo maior eficiência, precisão e agilidade na execução de diversas atividades.

### **2.1 conceito de inteligência artificial e aprendizado de máquina**

De acordo com Russel e Norvig (2004), a ciência da IA consiste em um campo universal, haja vista ser potencialmente relevante para qualquer esfera da atividade intelectual humana, declarando a importância do seu desenvolvimento e domínio na contemporaneidade.

Nas palavras dos sobreditos autores, pode-se afirmar que o primeiro trabalho de relevância reconhecido como IA foi realizado por Warren McCulloch e Walter Pitts, no ano de 1943, no qual esses pesquisadores sugeriram um modelo de neurônios artificiais, no qual cada neurônio desenvolvido era caracterizado por “ativo” e “inativo”, imitando o cérebro humano à luz da perspectiva da teoria da computação de Turing (Russel; Norvig, 2004).

Dentre as espécies que têm ganhado notória popularidade entre os indivíduos, destaca-se aquela conhecida como aprendizado de máquina (*Machine Learning*), a qual tem consistido em uma subárea essencial que possibilita que sistemas de IA aprendam a partir de grandes quantidades de dados, reconhecendo padrões e fazendo previsões sem terem que ser explicitamente programados para cada tarefa específica (Górriz, 2020).

Referido sistema de aprendizado de máquina é dividido em várias abordagens, como o aprendizado supervisionado, onde o algoritmo aprende a

partir de dados rotulados; o aprendizado não supervisionado, que busca padrões em dados não rotulados; e o aprendizado por reforço, onde o sistema aprende a tomar decisões através de recompensas e punições em um ambiente dinâmico (Górriz, 2020).

Tal sistemática dá origem a uma forma de utilizar as máquinas (*softwares* e *hardwares*) como ferramentas de redes neurais artificiais, inspiradas na estrutura e no funcionamento dos neurônios humanos. Essas redes são compostas por camadas de neurônios artificiais, denominados "nós", que processam informações e aprendem representações complexas de dados ao passá-las por várias camadas de processamento, um método conhecido como "*deep learning*" ou aprendizado profundo (Coutinho, Fernandes, 2019).

Tecnologias avançadas de IA, como redes neurais convolucionais (CNNs), usadas para reconhecimento de imagens, e redes neurais recorrentes (RNNs), empregadas no processamento de linguagem natural e análise de séries temporais, utilizam esses sistemas de aprendizado profundo, a propósito, sendo amplamente aplicado no âmbito do Direito, por sua capacidade de analisar grandes quantidades de dados jurídicos, prever resultados de litígios e automatizar procedimentos repetitivos (Queiroz, Gonçalves, 2024).

Outras formas de IA incluem as redes generativas adversariais (GANs), que consistem em duas redes neurais competindo entre si para criar dados sintéticos altamente realistas, além do processamento de linguagem natural (NLP), que permite que máquinas compreendam e gerem texto em linguagem humana (Stemler, 2023).

A complexa estruturação de tais redes possibilitam simular processos de pensamento humano e realizar tarefas que variam de simples automação até a tomada de decisões complexas, adaptando-se continuamente com base em novos dados e aprendendo de maneira autônoma para melhorar seu desempenho (Stemler, 2023).

Nesse contexto, o ChatGPT (sigla para *Generative Pre-Trained Transformer*) representa um exemplo notável de modelo de linguagem baseado em *deep learning* (aprendizagem profunda). A plataforma utiliza um algoritmo baseado em redes neurais que permite estabelecer uma conversa com o usuário por meio do processamento de um imenso volume de dados, se apoiando em milhares de exemplos de linguagem humana, para atender com profundidade o

contexto das solicitações dos usuários.

Assim, a contemporaneidade da IA a posiciona como uma ferramenta essencial para diversos setores, sobretudo tendo em vista que a IA pode absorver grandes quantidades de informação e deduzir resultados sem intervenção humana, contribuindo para avanços expressivos na produtividade e inovação tecnológica.

## **2.2. Aplicações da IA no Campo Jurídico**

A aplicação da IA no âmbito jurídico tem abrangido desde atividades como de análise de contratos e revisão de documentos até a previsão de resultados de litígios e suporte à decisão judicial, sob a promessa de redução de custos e aumento da eficiência de processos que antes dependiam exclusivamente de trabalho manual (Maia filho, 2018).

Perante essa perspectiva, pode-se afirmar principais aplicações da IA no Direito podem ser agrupadas em algumas categorias chave, sendo estas a análise e revisão de documentos, análise preditiva de litígios, suporte à decisão judicial, automação de tarefas jurídicas, e serviços de aconselhamento legal (Silva, 2023).

Nesse contexto, dentre as aplicações mais relevantes e amplamente adotadas da IA no Direito, cumpre dar enfoque em sua capacidade de catalisar a análise e revisão de contratos e documentos legais, eis que, mediante algoritmos de processamento de linguagem natural (NLP), os sistemas de IA são treinados para ler, interpretar e revisar documentos de forma rápida e precisa, identificando cláusulas específicas, verificando sua conformidade com normas legais e detectando padrões de jurisprudência em grandes volumes de dados (Elias, 2024).

Ademais, outra aplicação importante no Direito é a análise preditiva de litígios, que utiliza algoritmos de aprendizado de máquina para prever os resultados prováveis de processos judiciais com base em precedentes judiciais. Essas ferramentas, desenvolvidas com a partir de bancos de dados jurídicos, são capazes de examinar padrões em decisões judiciais anteriores, perfis de julgadores, tipos de litígios e diversas outras variáveis, oferecendo estimativas, por exemplo, quanto às chances de sucesso de um caso específico (Cueva,

2021).

Tais características fazem essa ferramenta ser extremamente valiosa para operadores do direito que pretendem formular de estratégias jurídicas, porquanto permitem uma avaliação de certa forma precisa dos riscos que envolvem o litígio, favorecendo a tomada de decisões com a intenção de alcançar desfechos processuais mais assertivos.

Softwares como o "*Ross Intelligence*", construído sobre o sistema de IA IBM Watson, exemplificam essa aplicação, ajudando operadores do direito a realizar pesquisas jurídicas mais rapidamente e fornecer respostas com base em enormes bases de dados de jurisprudências, além de estatutos e regulamentos (Paiva, 2022).

A IA também tem potencial significativo no suporte à decisão judicial, um campo que está ganhando atenção crescente devido à promessa de aumentar a imparcialidade e a eficiência das decisões legais. Nessa toada, Sebastião Tavares Pereira defende que a tecnologia permitirá a qualquer julgador "herculizar-se" possibilitando-o a:

considerar imparcialmente todos os sinais característicos relevantes de uma situação, numa dimensão de tempo adequada, à luz dos grandes princípios do ordenamento jurídico, de modo coerente e indutor da integridade do sistema" (Tavares Pereira, 2009, p.62).

Sistemas de suporte à decisão baseados em IA podem auxiliar juízes ao fornecer análises detalhadas de precedentes legais, identificar jurisprudências relevantes e até sugerir sentenças com base em casos semelhantes, possivelmente permitindo que julgadores decidam de uma forma mais legítima e paradigmática de prestação jurisdicional (Xu, 2022).

A propósito, em alguns sistemas judiciais, como na Estônia, já estão sendo experimentadas soluções de IA para a resolução de disputas de pequeno porte. Nessa experiência, algoritmos têm ajudado a determinar vereditos em questões relativamente simples, em que pese tenha feito surgir questionamentos quanto à transparência dos algoritmos, o devido processo legal e a possível erosão da autonomia judicial (Ribeiro, 2020).

Além disso, a IA está sendo cada vez mais utilizada na automação de tarefas jurídicas repetitivas e administrativas, como pesquisa jurídica, gestão de documentos, agendamento de audiências e controle de prazos processuais, corroborando com a seguinte possibilidade exposta por Fabiano Peixoto e

Roberta da Silva (2019. p. 119):

Uma segunda possibilidade é a IA como forma de apoio à demanda colocada ao Judiciário de celeridade e qualidade na prestação jurisdicional, inclusive como um mecanismo de *self optimization concept* dos servidores –extremamente qualificados, mas sendo utilizados nas atividades repetitivas e pouco estratégicas existentes pelo volume sem igual de processos judiciais.

A título de exemplo de aplicação dessas automações, a utilização de ferramentas de IA, como assistentes virtuais e *chatbots*, torna possível responder de maneira célere perguntas frequentes dos clientes e partes envolvidas em litígios e realizar a triagem inicial de conflitos processuais, oportunizando que trabalhadores do setor público e privado se concentrem em atividades mais estratégicas e analíticas (Nascimento, 2023).

Dessa forma, verifica-se que a IA está remodelando o panorama jurídico, tanto no ambiente público quanto privado, proporcionando maior eficiência e acessibilidade, ao passo que levanta desafios éticos e regulatórios importantes.

### **2.3 Vantagens da Utilização da IA no Direito**

Entre os principais benefícios da utilização da IA no sistema jurídico, destacam-se o aumento da celeridade processual e a redução de custos operacionais, não apenas acelerando a prática jurídica, mas também trazendo implicações profundamente benéficas para o acesso à justiça (Takaki, 2022).

Nessa senda, a IA, especialmente através de algoritmos de aprendizado de máquina e NLP, tem possibilitado automatizar grande parte das tarefas rotineiras dos operadores do direito, permitindo que sejam concluídas em uma fração do tempo que levariam se fossem realizadas manualmente (Dragalzev, 2024).

Ao automatizar tarefas processuais, como a triagem inicial de casos, a gestão de fluxos processuais, o agendamento de audiências e a elaboração automatizada de despachos, é certo que a IA pode garantir maior agilidade nas etapas processuais, ajudando a acelerar significativamente o tramite processual e a resolução de procedimentos judiciais.

Neste ponto, observa-se um grande benefício para o funcionamento

do sistema jurídico, sobretudo tendo em vista que a morosidade processual é um dos problemas mais persistentes e desafiadores enfrentados pelo sistema judicial brasileiro (Bastos, Da Fonsêca, 2011). Em consonância, disserta Barbara Brene que (2021, p. 9):

(...) a Inteligência Artificial tem sido vislumbrada como sendo um recurso promissor ao intento de solucionar o problema da morosidade processual, tendo por consideração, precipuamente, que a tecnificada do processo e da ciência do Direito no Brasil implica em diversas análises objetivas, tais como os requisitos de processabilidade de uma ação ou de um recurso, operações que facilmente podem ser operadas por meio de tecnologias, sem que isso implique em discussões quanto a imparcialidade ou precisão da máquina, uma vez que, ainda quando realizada pelo ser humana, trata-se de uma atividade meramente objetiva.

Ademais, outro benefício importante da utilização da IA no Direito é a redução de custos operacionais, haja vista que a implementação de sistemas de IA pode resultar em economias substanciais ao reduzir a necessidade de trabalho manual intensivo e ao minimizar os erros que podem gerar custos adicionais, como retrabalho e litígios desnecessários (Jacobsen, 2023).

Por exemplo, ao fazer com que a análise de grandes volumes de documentos seja automatizada, um processo que poderia exigir o trabalho de uma equipe inteira de operadores do direito pode ser realizado por um *software* de IA em um curto espaço de tempo, liberando recursos humanos e financeiros para outras atividades estratégicas.

Isso não significa, contudo, que a incorporação da IA no Direito implica a substituição dos operadores jurídicos, mas sim a sua potencialização, de modo a aprimorar eficiência e a qualidade do trabalho jurídico, auxiliando na pesquisa, na organização de informações e na automatização de tarefas repetitivas. Nessa senda, Tiago de Oliveira Siqueira bem destaca que (2020, p. 91):

Além de auxiliar e até orientar os profissionais, a Inteligência Artificial também pode gerar grandes revoluções e até criar mercados e modificar a forma com que vemos o mundo do direito (...) Não buscamos, contudo, tentar estabelecer um debate ético, pois em nenhum momento se pretende a substituição dos profissionais do direito, mas sim, a implementação de ferramentas que possam ampliar as possibilidades para um trabalho mais produtivo e benéfico à toda a sociedade.

Assim, depreende-se que as ferramentas baseadas em IA podem ajudar organizações de assistência jurídica a gerenciar melhor seus recursos limitados, direcionando os esforços dos operadores do Direito para os casos que

mais precisam de intervenção humana (Jacobsen, 2023).

À vista disso, verifica-se também que a democratização do acesso à justiça é outro benefício relevante proporcionado pela IA no Direito, pois ferramentas de IA, como plataformas de aconselhamento jurídico automatizado e triagem de casos, podem oferecer orientações preliminares a pessoas que não têm condições financeiras de contratar um advogado, permitindo que elas compreendam melhor seus direitos e as opções legais disponíveis (Bakshi, 2023).

Dessa forma, é certo que a aplicabilidade da IA no Direito tem trazido benefícios significativos para a área, influenciando a forma como os profissionais da área interagem com dados, legislações e sistemas jurídicos, malgrado levante debates sobre sua compatibilidade com os princípios fundamentais que rodeiam o campo jurídico, bem como sobre os limites e critérios para sua utilização nas atividades empenhadas.

## **2.4 Limitações e Problemas Inerentes à Implementação da IA no Direito**

Apesar das inúmeras vantagens proporcionadas pela implementação da IA no Direito, existem também limitações e problemas significativos que devem ser considerados de maneira crítica.

Posto isso, frisa-se que uma das limitações mais significativas da IA no Direito é a falta de compreensão contextual por parte dos algoritmos, uma vez que, embora a IA seja capaz de processar grandes volumes de dados e identificar padrões com uma precisão que supera a capacidade humana, ela ainda é fundamentalmente limitada no que concerne à compreensão de um contexto jurídico mais amplo, outrossim na consideração das nuances específicas de cada caso, de modo que suas análises podem ser errôneas a depender da situação. Corroborando essa intelecção, González bem destaca que (2023, p. 101):

Con ello se puede determinar que tanto la inteligencia artificial y el razonamiento humano realizan la misma función, pero no la misma efectividad, recordemos que una máquina y un ser humano no pueden ser comparados como tal, lo que sí se puede establecer es que el ordenador o computador realice una tarea similar a la que ejerce una persona.

Essa problemática se intensifica quando se dá enfoque no fato de que a prática jurídica muitas vezes envolve a interpretação de princípios legais complexos, a consideração de circunstâncias atenuantes ou agravantes e a aplicação de valores sociais e éticos (Oliani, 2019), emergindo o questionamento de que tais fatores podem não ser satisfatoriamente replicados por uma IA.

Nesse contexto, um algoritmo de IA pode, por exemplo, ser capaz de prever o resultado de um litígio com base em dados históricos, mas pode falhar ao considerar elementos subjetivos que influenciam uma decisão judicial, como a credibilidade das testemunhas ou a intenção moral das partes envolvidas (Magalhães, 2024).

Ademais, outra limitação importante é o problema da opacidade dos algoritmos, frequentemente referido como o problema da "caixa-preta". Muitos sistemas de IA, especialmente aqueles baseados em aprendizado profundo (*deep learning*), operam de maneiras que são opacas até mesmo para seus próprios criadores e usuários finais, fazendo com que, embora os sistemas possam fornecer resultados precisos, o processo exato pelo qual eles chegaram a esses resultados muitas vezes não é compreensível ou explicável (Alves, Andrade, 2022).

Essa falta de transparência é problemática sobretudo porque o princípio do devido processo legal exige que as decisões judiciais sejam transparentes e passíveis de revisão, de tal modo que, se uma decisão for influenciada ou tomada por um algoritmo cuja lógica não pode ser explicada, os direitos das partes ao devido processo e à defesa podem ser comprometidos, conforme se verifica das seguintes considerações de Nunes e Marques (2018, p. 7):

A ausência de transparência do algoritmo é especialmente crítica nesse caso. Como defender-se de um "índice" sem saber o método de seu cálculo? Como submeter o "índice" ao controle do devido processo constitucional? Por mais que sejam divulgadas as perguntas realizadas, os acusados não sabem como suas respostas influenciam no resultado final (output). Dessa forma, a defesa do acusado torna-se impossibilitada por dados matemáticos opacos e algoritmicamente enviesados, mas camuflados, pela "segurança" da matemática, como supostamente imparciais, impessoais e justos.

Além disso, a possibilidade de perpetuação e amplificação de vieses é outro problema crítico que surge com a implementação de IA no Direito, haja vista que os algoritmos de IA são tão imparciais quanto os dados em que são

treinados. Dessa forma, se os dados de treinamento contêm vieses históricos, preconceitos ou discriminações, o algoritmo inevitavelmente aprenderá e reproduzirá esses vieses em suas previsões e decisões, sendo, claramente, uma possível problemática de aplicação da IA na elaboração automatizada de decisões judiciais (Ferrara, 2023).

Por exemplo, um algoritmo treinado em dados judiciais históricos pode internalizar preconceitos raciais ou de gênero presentes nos dados e replicá-los nas previsões de risco de reincidência ou na recomendação de penas, passando a inadvertidamente reforçar desigualdades estruturais se esses vieses não forem identificados e corrigidos. Nessa toada, Ferreira Filho adverte que (2023, p. 16):

A tecnologia, quando usada sem sabedoria, pode ameaçar a dignidade da pessoa humana, transformando o Direito em uma ciência exata, onde as particularidades e nuances de cada caso são ignoradas. A busca por eficiência no judiciário não deve sacrificar a humanidade e a justiça que o sistema deve oferecer. O avanço tecnológico deve ser acompanhado por uma análise crítica de seus impactos, tanto positivos quanto negativos, e medidas adequadas para garantir que a dignidade da pessoa humana seja preservada.

Por consequência, a atribuição de responsabilidade legal torna-se outra questão complexa associada ao uso da IA no Direito, visto que, quando um erro é cometido por um sistema de IA – seja uma decisão jurídica errada, uma recomendação enviesada ou uma violação de direitos – torna-se difícil determinar quem deve ser responsabilizado, surgindo a dúvida se a responsabilidade sobre o desenvolvedor do algoritmo, a empresa que fornece a tecnologia, o operador que utiliza a ferramenta ou a instituição que decide implementar a IA (Nersessian, 2020).

Assim, a dispersão de responsabilidade e a complexidade técnica dos sistemas de IA podem dificultar a atribuição clara de culpabilidade e a busca de reparação por danos, criando um vácuo legal que precisa ser preenchido com regulamentações específicas que definam critérios de responsabilidade e prestação de contas (Nersessian, 2020).

Além desses problemas, existem também desafios técnicos e práticos que precisam ser considerados, como o fato de que a precisão dos sistemas de IA depende fortemente da qualidade dos dados que alimentam esses algoritmos. Dados desatualizados, incompletos ou mal interpretados podem levar a erros

significativos, afetando diretamente a qualidade das decisões jurídicas automatizadas (Marco; Lemes, 2020).

Por fim, é importante considerar a dependência excessiva de decisões automatizadas, que pode levar à "desumanização" da prática jurídica, pois a confiança excessiva em sistemas de IA pode levar a uma abordagem mecanicista da justiça, onde o valor humano e o discernimento são minimizados. tornando ideal que, embora a IA possa ser uma ferramenta poderosa para auxiliar na prática jurídica, é crucial que ela seja utilizada como um complemento ao julgamento humano, e não como um substituto. (Ferreira Filho, 2023).

Tais problemáticas emergem desafios éticos quanto à aplicação da IA no Direito, especialmente no que se refere à transparência, imparcialidade e responsabilidade, levantando questões sobre como garantir um processo justo e equitativo em um contexto altamente influenciado por tecnologias de IA.

### **3. DESAFIOS ÉTICOS DA INTELIGÊNCIA ARTIFICIAL NO SISTEMA JURÍDICO**

A imparcialidade e transparência são pilares fundamentais do sistema de Justiça e do Estado de Direito. No entanto, a crescente utilização de IA no campo jurídico trouxe à tona preocupações significativas sobre a imparcialidade e a discriminação algorítmica, sobretudo considerando a necessária preservação das garantias individuais nos Estados Democráticos de Direito.

#### **3.1 Imparcialidade e Discriminação Algorítmica**

Algoritmos, especialmente aqueles utilizados em sistemas de IA e aprendizado de máquina, têm o potencial de reproduzir e até amplificar preconceitos existentes nos dados em que são treinados, resultando em discriminação sistêmica e, por consequência, levantando questões críticas sobre como a tecnologia deve ser regulamentada e supervisionada para garantir a promoção de um sistema legal justo e imparcial (Ferrara, 2023).

A propósito, cunhou-se o termo "viés algorítmico" para se referir à distorção sistemática nos resultados de um algoritmo que leva a decisões ou previsões tendenciosas. Esses vieses podem ser introduzidos em diferentes etapas do desenvolvimento e implementação de um algoritmo, como durante o processo de coleta de dados, fazendo com que dados lastreados em preconceitos históricos sejam utilizados para instruir o treinamento do sistema (Carvalho, 2020).

Portanto, se esses dados não forem cuidadosamente analisados e tratados, o algoritmo aprenderá e replicará essas aberrações, podendo resultar em decisões distorcidas e enviesadas. No contexto jurídico, isso é especialmente problemático, pois a justiça exige que as decisões sejam baseadas em princípios legais e evidências objetivas, não em preconceitos subjacentes, tendo em vista que somente através dessa metodologia se conserva um Estado Democrático de Direito (Ferrajoli, 2002).

Um exemplo claro de como vieses algorítmicos podem levar à discriminação sistêmica é o uso de sistemas de IA para prever o risco de reincidência em processos criminais, como o sistema COMPAS (*Correctional*

*Offender Management Profiling for Alternative Sanctions*) utilizado nos Estados Unidos (Alves, Pimentel, 2022).

Estudos demonstraram que o COMPAS tendia a superestimar o risco de reincidência de réus negros, enquanto subestimava o risco para réus brancos. Isso corria porque os algoritmos foram treinados em dados históricos de justiça criminal que já estavam contaminados por preconceitos raciais, resultando em decisões judiciais injustas e na perpetuação de um ciclo de discriminação racial dentro do sistema de justiça (Alves, Pimentel, 2022).

Assim, para abordar o problema da imparcialidade e discriminação algorítmica, é crucial adotar uma abordagem multifacetada que envolva tanto a regulamentação quanto a supervisão ética, devendo-se, em primeiro lugar, haver um foco na auditoria e validação de dados utilizados para treinar algoritmos de IA (Leal, 2023).

Desse modo, dados históricos e de treinamento são cuidadosamente examinados e corrigidos para remover vieses que possam influenciar negativamente as decisões algorítmicas, garantindo a idoneidade das conclusões adotadas pela IA (Leal, 2023). Isso pode incluir a utilização de técnicas de aprendizado de máquina justas (*fair machine learning*), que buscam minimizar o impacto de vieses existentes e garantir que as decisões sejam feitas com base em critérios justos e objetivos (Grgić; Zafar; Gummadi; Weller, 2018).

Não se pode olvidar, ainda, a necessidade de incorporação de diretrizes éticas e princípios de responsabilidade no desenvolvimento e implementação de sistemas de IA no Direito, de modo que desenvolvedores de IA, operadores jurídicos e formuladores de políticas devem trabalhar juntos para criar normas que assegurem que a tecnologia seja usada de maneira adequada, garantindo um sistema jurídico mais equitativo, sem comprometer os direitos fundamentais e a integridade processual (Carvalho, 2020).

### **3.2. Transparência e o Problema da "Caixa-Preta"**

A transparência é um princípio fundamental para o funcionamento adequado do sistema jurídico, uma vez que garante o direito ao devido processo legal, promove a justiça e mantém a confiança pública nas instituições (Warmling; Bernardes; Santos, 2014).

No entanto, a crescente adoção de sistemas de IA no Direito introduz o problema da "caixa-preta" na sua utilização, um desafio significativo relacionado à falta de explicabilidade dos algoritmos que tomam ou influenciam decisões jurídicas (Alves; Andrade, 2022).

O termo "caixa-preta" refere-se à opacidade dos processos de tomada de decisão em muitos sistemas de IA, particularmente aqueles baseados em aprendizado profundo, onde as regras e critérios usados para alcançar um resultado não são facilmente compreensíveis ou explicáveis, nem mesmo para os desenvolvedores do algoritmo (Valença, 2007).

Essa falta de transparência pode minar princípios jurídicos fundamentais, como a responsabilização, a contestabilidade e o acesso à justiça, impactando diretamente o direito ao devido processo legal e a confiança pública no sistema de justiça (Maranhão; Junquilha; Tasso, 2023).

O problema da "caixa-preta" dos algoritmos de IA surge principalmente devido à complexidade e ao caráter não-linear dos modelos de aprendizado profundo, que muitas vezes envolvem milhões de parâmetros ajustados através de camadas múltiplas de neurônios artificiais. Embora esses modelos possam oferecer previsões precisas e desempenhos superiores em várias tarefas, como a análise preditiva de litígios ou a revisão de contratos, a lógica subjacente que orienta essas previsões é, na maioria das vezes, ininteligível para seres humanos, impedindo uma compreensão exata do processo utilizado para o resultado apresentado (Gripon, 2020).

Diferentemente de um programa tradicional, onde cada passo é codificado explicitamente e pode ser revisado, os sistemas de aprendizado profundo aprendem padrões a partir de grandes volumes de dados, desenvolvendo internamente regras que não são facilmente acessíveis ou compreensíveis. Isso resulta em uma falta de explicabilidade, ou "explicabilidade algorítmica", que é justamente a capacidade de um modelo de IA de fornecer uma justificativa clara e compreensível para suas decisões ou previsões. (Gripon, 2020)

À vista disso, cumpre ressaltar que a falta de explicabilidade é particularmente problemática no âmbito jurídico, sobretudo porque, em um Estado Democrático de Direito, o devido processo legal garante que as partes envolvidas em um litígio tenham o direito de entender como e por que uma

decisão foi tomada, a fim de que possam contestá-la de maneira justa e adequada (Fioratto; Dias, 2010).

Nessa mesma inteligência, destaca-se que o princípio do devido processo exige transparência não apenas na aplicação da lei, mas também na explicação de como as decisões são alcançadas. Sem essa transparência, o sistema jurídico corre o risco de se tornar arbitrário, minando a confiança pública nas instituições judiciais e violando direitos fundamentais (Fioratto; Dias, 2010).

Assim, se uma decisão é influenciada ou tomada por um algoritmo de IA cuja lógica é opaca, as partes afetadas podem não ter meios de compreender a base da decisão, o que pode refletir em uma situação de que as decisões legais se tornam inacessíveis e imunes ao escrutínio público, comprometendo a justiça processual e o direito das partes à defesa (Rudin; Wang; Coker, 2020).

Além disso, a própria falta de transparência na forma de processamento dos algoritmos de IA pode prejudicar a confiança pública no sistema de justiça, uma vez que essa opacidade pode dar a falsa impressão de que decisões automatizadas são inquestionavelmente corretas, quando na realidade, considera-se que elas podem conter vieses ocultos ou ser baseadas em dados inadequados (Rudin; Wang; Coker, 2020).

Assim, verifica-se que para mitigar os riscos associados ao problema da "caixa-preta", é necessário promover uma maior transparência e explicabilidade nos sistemas de IA aplicados ao Direito. Isso pode envolver a implementação de metodologias de IA explicável (Explainable AI - XAI), que se refere ao desenvolvimento de modelos de IA que não apenas forneçam resultados precisos, mas também sejam capazes de explicar suas decisões de maneira compreensível para os humanos (Alves; Andrade, 2022).

Ferramentas como LIME (*Local Interpretable Model-agnostic Explanations*) e SHAP (*SHapley Additive exPlanations*) têm sido desenvolvidas para ajudar a tornar os modelos de IA mais interpretáveis, permitindo que os usuários compreendam quais características ou dados influenciaram uma determinada decisão (Aditya, 2022).

No entanto, enquanto essas ferramentas representam um avanço importante, a aplicação de XAI no contexto jurídico ainda enfrenta desafios, uma vez que o nível de explicabilidade exigido no Direito pode ser mais rigoroso do que em outras áreas, dada a necessidade de justificar decisões de maneira clara

e convincente (Deeks, 2019).

### **3.3. Responsabilidade Legal e Ética no Uso da IA**

Além do problema de falta de transparência no trabalho dos algoritmos, a crescente adoção da IA no Direito levanta questões complexas sobre a responsabilidade legal e ética quando decisões automatizadas resultam em resultados injustos ou prejudiciais. Essa problemática se intensifica ante o fato da responsabilidade legal em relação ao uso da IA ser multifacetada e envolver a análise de aspectos de responsabilidade civil e penal, assim como questões éticas relacionadas à transparência, justiça e proteção dos direitos humanos (Nersessian, 2020).

A responsabilidade civil é uma das áreas centrais de discussão quando se trata de decisões de IA que levam a resultados injustos ou prejudiciais. No Direito Civil, a responsabilidade por danos geralmente recai sobre a parte que, por ação ou omissão, causa um prejuízo a outra parte (Fonseca, 2017).

Em contrapartida, quando se trata de decisões automatizadas por IA, a questão de quem deve ser responsabilizado por um erro é menos clara e envolve múltiplos atores: o desenvolvedor do algoritmo, o fornecedor da tecnologia, o operador (como um advogado ou um tribunal) e a instituição que implementa a IA (como um escritório de advocacia ou o próprio sistema de justiça) (Mcdonald, 2023).

Nessa toada, destaca-se ainda que a complexidade técnica dos sistemas de IA e a falta de transparência sobre como os algoritmos tomam decisões complicam ainda mais a atribuição de responsabilidade, gerando incerteza sobre quem deve ser responsabilizado por prejuízos resultantes das decisões automatizadas (Reed, 2018).

Por exemplo, se um sistema de IA usado para recomendar sentenças judiciais sugere uma punição desproporcionalmente severa com base em dados enviesados, quem deve ser responsabilizado? O desenvolvedor que criou o algoritmo, a instituição que o utilizou sem a devida supervisão, ou o próprio operador que confiou cegamente nas recomendações da IA?

Via de regra, a culpa pode ser atribuída a um agente humano específico que causou o dano por negligência ou ação intencional. No entanto,

no caso da IA, o agente que causa o dano é um algoritmo complexo que opera com um grau de autonomia, e a responsabilidade pode estar fragmentada entre várias partes, evidenciando que, em muitos casos, o problema reside na distribuição da responsabilidade (Benhamou, 2020).

Uma possível abordagem para lidar com essa fragmentação é a introdução do conceito de "responsabilidade solidária", onde todos os agentes envolvidos no desenvolvimento, implementação e uso da IA poderiam ser considerados responsáveis em diferentes graus. Essa abordagem, no entanto, pode levar a problemas de incerteza jurídica e sobrecarga para os operadores de IA, o que poderia desencorajar a inovação e o uso da IA no Direito (McDonald, 2023).

Ademais, além da responsabilidade civil, frisa-se que a responsabilidade penal levanta problemáticas ainda mais maiores quando se trata do uso de IA no Direito, eis que o Direito Penal se baseia na premissa de que a responsabilidade criminal exige a existência de dolo (intenção) ou culpa (negligência, imprudência ou imperícia) (Souza, 2010).

Nesse sentido, ressalta-se que a IA, como uma ferramenta de decisão automatizada, não possui intenção ou consciência e, portanto, não pode ser considerada criminalmente responsável, reforçando o desafio de identificar qual agente humano ou entidade corporativa pode ser considerado culpado quando uma decisão de IA resulta em um resultado penalmente relevante (García, 2021).

Uma abordagem poderia ser responsabilizar o desenvolvedor da IA ou a empresa que forneceu o sistema, especialmente se o software foi desenvolvido de forma negligente, sem as salvaguardas necessárias para evitar danos. No entanto, provar a negligência ou o dolo em tais casos pode ser extremamente complicado, especialmente quando os algoritmos operam como "caixas-pretas", dificultando a análise de onde exatamente o erro ocorreu (Martin, 2019).

Outra perspectiva importante no debate sobre responsabilidade legal e ética no uso da IA é a questão do "dever de cuidado". No uso de IA no contexto jurídico, deve haver um dever de cuidado explícito por parte daqueles que desenvolvem, implementam e utilizam sistemas de IA para garantir que eles sejam projetados, treinados e operados de maneira justa e segura (Bashayreh, 2023).

Isso inclui a obrigação de realizar auditorias regulares de vieses nos dados de treinamento, garantir que os algoritmos sejam transparentes e explicáveis, e fornecer instruções claras aos operadores sobre como interpretar e utilizar os resultados da IA. A falha em cumprir com esses deveres de cuidado pode constituir uma base para responsabilidade legal, tanto em termos civis quanto penais, dependendo da gravidade do dano causado e do grau de negligência envolvido.

Do ponto de vista ético, a responsabilidade no uso da IA envolve mais do que apenas a atribuição de culpa legal, requerendo uma reflexão sobre os princípios fundamentais que devem orientar o desenvolvimento e a aplicação da IA no Direito.

A utilização de IA deve estar alinhada com os princípios éticos de justiça, equidade, transparência e respeito aos direitos humanos, de tal forma que desenvolvedores de IA, operadores jurídicos e formuladores de políticas têm uma responsabilidade ética de garantir que a IA seja usada para promover o bem-estar social e evitar a perpetuação de injustiças, incluindo mecanismos que assegurem que os algoritmos sejam projetados para serem justos, que as decisões automatizadas sejam explicáveis e que haja um sistema claro de prestação de contas quando ocorrerem erros ou injustiças (Cath, 2018).

No que concerne especificamente à regulamentação, vários países e organizações internacionais estão começando a explorar novas formas de responsabilização para IA, propondo marcos legais que reconheçam a complexidade e a especificidade desses sistemas. A União Europeia, por exemplo, através da proposta de Regulação de IA, está considerando um modelo de responsabilidade baseado em risco, onde a responsabilidade é proporcional ao nível de risco associado ao uso de IA em contextos específicos (Paul, 2023).

Por fim, uma abordagem alternativa discutida na literatura é a criação de um regime de "personalidade eletrônica" para sistemas de IA. Isso significaria que a IA, em certos contextos, poderia ser tratada como uma entidade jurídica separada, com direitos e responsabilidades próprias, semelhante à maneira como as corporações têm personalidade jurídica (Teubner, 2018).

Conforme abordado por Teubner (2018), o conceito de personalidade eletrônica permitiria que sistemas de IA autônomos pudessem ser tratados como "agentes" que podem ser responsabilizados por suas decisões e ações no

âmbito jurídico, oportunizando uma nova fronteira na responsabilização desses sistemas.

Em que pese essa ideia seja controversa e tenha muitas implicações filosóficas e legais, ela reflete a necessidade de pensar de forma inovadora sobre a responsabilidade na era da IA, cabendo ressaltar ainda que a atribuição de personalidade eletrônica a sistemas de IA também levanta questões sobre a natureza da consciência e da agência, já que a IA, ao contrário dos seres humanos e das corporações, não possui a capacidade de intencionalidade ou moralidade, de modo a exigir uma profunda reavaliação de como a sociedade entende a "responsabilidade", tanto em termos de direito quanto de ética (Ianni; Monterossi, 2017).

### **3.4. Privacidade e Proteção de Dados**

Como exposto ao longo deste trabalho, a IA, por sua natureza, depende de grandes volumes de dados para aprender e melhorar seu desempenho em diversas tarefas, incluindo aquelas aplicadas ao Direito.

Não obstante, o uso extensivo de dados, mormente dados pessoais, levanta sérios desafios em relação à privacidade e à proteção de dados, os quais são amplificados no âmbito do Direito, onde a confidencialidade e a integridade dos dados são essenciais para assegurar os direitos fundamentais dos indivíduos e a confiança no sistema judiciário (McDonald, 2023).

Assim, ao passo que sistemas de IA são cada vez mais utilizados para análise preditiva de litígios, revisão de documentos legais e até para a tomada de decisões judiciais, surge a necessidade urgente de considerar como esses sistemas devem ser regulamentados para proteger a privacidade e os dados pessoais, a fim de assegurar que essa implementação seja feita de maneira ética e legalmente responsável (Teubner, 2018).

Nessa toada, destaca-se que o desafio da proteção de dados pessoais no uso de IA no Direito começa com a própria natureza dos dados utilizados, uma vez que, no campo jurídico, as etapas de processamento massivo de dados frequentemente envolvem o uso de dados altamente sensíveis, como informações pessoais de partes envolvidas em litígios, antecedentes criminais, dados financeiros e históricos médicos. Dessa forma,

quando manipulados sem a devida proteção, os cidadãos podem ser vulneráveis a violações de privacidade, como acessos não autorizados, vazamentos ou até usos indevidos (Iani; Monterossi, 2017).

À vista desse imbróglio, há de se destacar que uma das principais preocupações relacionadas ao uso de IA no Direito é o risco de reidentificação de dados, pois, mesmo quando os dados são anonimizados ou pseudonimizados, os algoritmos de IA podem, com frequência, recombina dados e identificar indivíduos com base em informações aparentemente inofensivas, dando azo a violações ao direito à privacidade e à consumação de consequências jurídicas e éticas severas (Doneda Et Al., 2018).

Por exemplo, informações de testemunhas protegidas ou detalhes de casos sensíveis poderiam ser inadvertidamente revelados, comprometendo não apenas o direito à privacidade, mas também a segurança das partes envolvidas.

Ademais, outro desafio significativo é a transferência transfronteiriça de dados, pois, em um mundo cada vez mais globalizado, sistemas de IA frequentemente envolvem o processamento de dados em diferentes jurisdições, cada uma com suas próprias leis e regulamentos de proteção de dados.

Citando um exemplo real dessa problemática, dados coletados em um país com rigorosas leis de proteção de dados, como o Regulamento Geral de Proteção de Dados (GDPR) da União Europeia, podem ser transferidos e processados em outra jurisdição com padrões de proteção de dados menos rígidos, criando um dilema sobre qual regime jurídico se aplica e como garantir que a privacidade dos dados seja mantida em todas as fases de processamento (Bielova; Byelov, 2023).

Há de se destacar que instrumentos legais, como o GDPR na União Europeia, já estabelecem requisitos rigorosos para a utilização de dados pessoais, incluindo princípios de minimização de dados, limitação de propósito, transparência e consentimento explícito (Brkan, 2017).

No entanto, apesar dessas regulamentações, ainda há lacunas significativas quanto a devida utilização desses dados, sobretudo considerando a evolução e a diversidade na utilização das IAs. Assim, mesmo com esses avanços regulatórios, a rápida evolução das tecnologias de IA desafia a capacidade de as normas acompanharem, deixando incertezas quanto à proteção adequada dos dados pessoais (Frazier, 2020).

Acrescenta-se que além da conformidade com as regulamentações legais, deve haver um compromisso ético por parte de desenvolvedores e operadores de IA para garantir que os dados sejam tratados com o máximo cuidado e que os princípios de privacidade sejam respeitados, fazendo surgir outra questão: a conformidade ética no uso de IA para proteger a privacidade e os dados pessoais (Martin; Rubinstein, 2020).

À vista disso, pode-se afirmar que a contemplação desse compromisso pode partir da adoção de práticas de "privacidade por design", onde a proteção de dados é integrada no desenvolvimento de sistemas de IA desde o início, sendo necessário, ainda, garantir que haja um controle contínuo sobre o uso e o armazenamento de dados para identificar e mitigar quaisquer riscos emergentes (Cavoukian, 2019).

Além disso, o desenvolvimento de protocolos de governança de dados específicos para IA no contexto jurídico pode ajudar a estabelecer diretrizes claras sobre como os dados devem ser coletados, armazenados, processados e compartilhados, a fim de garantir que a privacidade seja protegida e que a responsabilidade seja devidamente atribuída em casos de violações de dados.

Posto isso, é assertivo afirmar que a previsão de auditorias e a supervisão contínua são particularmente importantes para garantir que os sistemas estejam em conformidade com as normativas de proteção de dados, pois permitem a identificação e correção de possíveis falhas ou vulnerabilidades que possam comprometer a integridade dos dados, garantindo, assim, maior transparência e segurança no uso de IA (Doneda et al., 2018).

### **3.5. Autonomia e Controle Humano nas Decisões Automatizadas**

Apesar de a IA oferecer benefícios significativos em termos de eficiência, celeridade e precisão, a substituição de julgamentos humanos por sistemas automatizados levanta questões éticas profundas e desafiadoras.

Nesse sentido, ressalta-se que a crescente dependência de sistemas automatizados ameaça deslocar essa autonomia, transferindo-a para algoritmos que, embora sofisticados, carecem de compreensão contextual e ética, impondo uma reflexão crítica sobre os limites éticos da automação e as salvaguardas

necessárias para garantir que a utilização de IA no Direito fortaleça a justiça e a equidade (Gasparian; Turko; Besschetnova, 2023).

Posto isso, dá-se enfoque no fato de que um dos riscos mais evidentes da redução da autonomia humana nas decisões automatizadas é a erosão do julgamento humano, que é essencial para a prática jurídica. No sistema de justiça, juízes, advogados e outros operadores do direito não apenas aplicam leis de forma mecânica; eles interpretam textos legais, consideram as circunstâncias únicas de cada caso e ponderam os princípios de justiça e equidade (Rogers; Bell, 2019).

Por exemplo, ao julgar um caso de direito de família, um juiz pode considerar o bem-estar emocional das partes envolvidas e tomar uma decisão que, embora não seja a solução mais "eficiente" ou "objetiva" segundo critérios de IA, seja a mais justa e humana.

Desse modo, a dependência excessiva de sistemas automatizados pode levar a uma "desumanização" da justiça, onde a singularidade de cada caso é substituída por padrões estatísticos e resultados preditivos, de forma que tais decisões baseadas em algoritmos de aprendizado de máquina pode, por exemplo, recomendar uma sentença com base em dados históricos de decisões anteriores, sem considerar os contextos específicos ou as mudanças nos valores sociais (Chistilina, 2021).

Por efeito, isso pode resultar em uma justiça rígida e insensível, que, em vez de promover a equidade, acaba perpetuando injustiças estruturais (Zuckerman, 2020). Frente a isso, surge o risco de que, ao confiar cegamente em decisões automatizadas, os operadores do direito possam deixar de exercer o julgamento crítico necessário para questionar, revisar e adaptar decisões às realidades dinâmicas e variadas da sociedade, tornando o sistema judiciário insensível e distante da complexidade humana (Bebiano, 2022).

Esse ponto faz surgir outra problemática: a redução da autonomia humana nas decisões automatizadas pode criar uma espécie de "autoridade algorítmica", onde as decisões dos sistemas de IA são vistas como objetivas e indiscutíveis (Martin, 2018).

Em suma, o conceito de autoridade algorítmica refere-se à tendência de aceitar as decisões dos algoritmos como superiores, devido à percepção de que são mais precisas e livres de vieses humanos. No entanto, os algoritmos

não são intrinsecamente neutros, sobretudo porque são criados e treinados por humanos que trazem seus próprios preconceitos e limitações (Martin, 2018).

Desse modo, quando as decisões de IA são tomadas como infalíveis, pode-se criar uma falsa sensação de segurança que desencoraja o questionamento crítico e a contestação, podendo comprometer a capacidade das partes de buscar justiça de maneira plena e justa.

Assim, para enfrentar esses desafios, é fundamental considerar os limites éticos da automação no Direito, devendo-se primeiramente reconhecer que a automação não pode e não deve substituir completamente o julgamento humano em áreas que envolvem direitos fundamentais, ética e valores sociais complexos (Cath, 2018).

Em vez disso, a IA deve ser usada como uma ferramenta de apoio à decisão que ajuda os operadores do direito como, por exemplo, para analisar grandes volumes de informações, identificar padrões e prever resultados possíveis, mas sempre sob a supervisão e o controle humano, de modo que a decisão final deve ser baseada em um equilíbrio entre os *insights* fornecidos pela IA e o julgamento humano (Martin, 2019).

Por outro lado, deve haver normas regulatórias claras e diretrizes éticas que estabeleçam os limites de como e quando a IA pode ser usada para tomar decisões jurídicas, as quais devem garantir que os operadores humanos mantenham um controle significativo sobre o processo de tomada de decisão e que a IA seja usada de maneira transparente, explicável e contestável (Engelmann, 2023).

Essa regulamentação deve incluir requisitos para a validação regular dos algoritmos, auditorias de vieses e a presença de mecanismos de contestação eficazes que permitam às partes afetadas questionar e revisar decisões automatizadas, sendo ideal para garantir a justiça e equidade no processo jurídico, outrossim a responsabilização pelos resultados gerados pelas decisões.

Finalmente, verifica-se do exposto que o uso de IA no Direito deve ser uma ferramenta subsidiária ante a ação humana, sugerindo que a IA deva ser utilizada apenas em situações em que o julgamento humano é incapaz de lidar com a complexidade da tarefa, ou onde a IA pode claramente fornecer suporte benéfico, garantindo que a IA não substitua a capacidade humana de tomar

decisões, mas sim a complemente para o alcance de um melhor resultado (Zhao, 2021).

## **4. A FILOSOFIA DE HANS JONAS COMO BALUARTE PARA A REGULAÇÃO DA IA NO DIREITO**

Hans Jonas, filósofo alemão do século XX, teve uma influência significativa na reflexão ética sobre a tecnologia, incluindo a IA. A aplicação do seu pensamento na utilização da IA moderna sugere a necessidade de regulamentação, transparência e um compromisso ético com o bem comum, convidando a sociedade a pensar criticamente sobre a relação entre humanidade e tecnologia, ao passo que enfatiza que a necessidade de responsabilidade ante as evoluções tecnológicas para garantir um futuro seguro e sustentável (Rodríguez, 2022).

### **4.1 A Ética da Responsabilidade de Hans Jonas**

A Ética da Responsabilidade de Hans Jonas, apresentada em sua obra "*O Princípio Responsabilidade: Ensaio de uma Ética para a Civilização Tecnológica*", propõe um novo paradigma ético adaptado às demandas e aos riscos da era tecnológica, caracterizada pela capacidade humana de transformar radicalmente o mundo e, potencialmente, colocar em risco a própria existência da humanidade e da natureza (Jonas, 2006).

Sua perspectiva parte do entendimento de que, em razão da crescente autonomia da técnica e da ciência, os seres humanos passaram a ter um poder de intervenção sobre o mundo natural e social sem precedentes. Esse poder, no entanto, carrega consigo responsabilidades inéditas, nunca antes previstas pela humanidade, de modo que suas consequências podem afetar não só as gerações atuais, mas também as futuras.

Nessa senda, Jonas argumenta que, diante dos poderes sem precedentes conferidos pela tecnologia, o pensamento ético tradicional, que se limitava às consequências imediatas e diretas das ações humanas, precisa ser revisitado para incorporar uma preocupação profunda com as gerações futuras e com o impacto de longo prazo das nossas decisões tecnológicas (Gonçalves; Torres, 2020). Conforme destaca o filósofo:

A presença do homem no mundo era um dado primário e indiscutível de onde partir toda idéia de dever referente à conduta humana: agora, ela própria tornou-se um objeto de dever – isto é, o dever de proteger

a premissa básica de todo o dever, ou seja, precisamente a presença de meros candidatos a um universo moral no mundo físico do futuro; isso significa, entre outras coisas, conservar este mundo físico de modo que as condições para uma tal presença permaneçam intactas; e isso significa proteger a sua vulnerabilidade diante de uma ameaça dessas condições. (Jonas, 2006, p. 45).

Assim, diferente das éticas tradicionais, centradas na moralidade do indivíduo e na reciprocidade imediata entre os sujeitos, a Ética da Responsabilidade introduz um novo paradigma que leva em consideração as consequências a longo prazo das ações humanas, incluindo seus impactos sobre as gerações futuras e sobre a natureza (Nodari, 2014).

#### **4.1.1 A necessidade da modernização da definição de ética**

As éticas clássicas, fundamentadas em princípios religiosos, filosóficos e morais, foram concebidas para contextos em que a ação humana possuía um impacto restrito, limitado ao tempo e ao espaço imediato. No entanto, de acordo com Jonas, a revolução tecnológica transformou radicalmente essa realidade, conferindo ao ser humano um poder sem precedentes sobre a natureza e sobre seu próprio destino.

Nesse contexto, o agir humano deixou de se restringir às interações interpessoais e passou a influenciar de maneira irreversível o meio ambiente e as futuras gerações, de modo que os desenvolvimentos contemporâneos fazem com que seus colaterais reverberem no tempo. Diante dessa nova configuração, Jonas argumenta que o princípio da responsabilidade deve ser a base de uma nova ética, que leve em consideração as consequências de longo prazo da ação tecnológica (Fonseca, 2012).

O filósofo alemão destaca que a tecnologia contemporânea não pode mais ser vista apenas como um instrumento neutro à disposição do homem. Pelo contrário, ela adquiriu uma dinâmica própria, em um processo contínuo de autoprocriação e autoprogressão.

Isso faz com que éticas tradicionais, como as formuladas por Aristóteles, Kant e os utilitaristas como Mill e Bentham, não são capazes de oferecer respostas satisfatórias para esse novo contexto, sobretudo tendo em vista que as novas tecnologias têm proporcionado uma inversão dos papéis entre

o ser humano e a técnica: o homem, antes agente e sujeito, passa a ser dominado pelas exigências do progresso tecnológico, que impõe seu próprio ritmo e objetivos. (Fonseca, 2012).

Por exemplo, o imperativo categórico kantiano, baseado na universalização da ação e na autonomia do sujeito moral, não prevê o alcance intergeracional das decisões tecnológicas atuais, ao passo que o utilitarismo, por sua vez, com sua busca pela maximização da felicidade, não consegue lidar adequadamente com os riscos incalculáveis inerentes ao progresso técnico, fazendo com que tais filosofias sintetizadas nos séculos passados sejam incapazes de oferecer respostas satisfatórias para esse novo contexto.

Assim, Jonas sintetiza o princípio ético da responsabilidade, o qual rompe com a separação tradicional entre ética e conhecimento científico, introduzindo um saber preditivo que antecipa e avalia os impactos das inovações tecnológicas antes de sua implementação (Fonseca, 2012).

#### **4.1.2 Fundamentos da Ética da Responsabilidade**

Jonas defende que o ideal baconiano, que propunha o uso do conhecimento para subjugar a natureza e melhorar a condição humana, acabou por se converter, no capitalismo moderno, em um sistema irracional, predatório e antiético, resultando em um desenvolvimento descontrolado, marcado por produção excessiva e consumo exacerbado, que ameaça tanto o equilíbrio social quanto a sustentabilidade ambiental.

Dessa forma, o imperativo da responsabilidade se estrutura de tal forma a impor limites ao desenvolvimento técnico-científico, de forma a evitar a catástrofe planetária e garantir a continuidade da vida humana e do meio ambiente. Nessa senda, a Ética da Responsabilidade de Jonas propõe uma reorientação das decisões humanas a partir de uma perspectiva temporal ampliada, tornando o futuro da humanidade o primeiro dever da conduta coletiva na era da civilização técnica (Barbosa, 2023).

Para garantir a sustentabilidade da vida, Jonas introduz o conceito da heurística do medo, que se contrapõe às promessas otimistas do progresso ilimitado. O medo, nesse contexto, não deve ser visto como uma paralisia diante do futuro, mas como um motor de ação e responsabilidade. Ele nos alerta para

os perigos inerentes ao poder tecnológico descontrolado e nos impele a adotar medidas prudentes para evitar catástrofes ambientais e sociais (Barbosa, 2023).

Com essa perspectiva, o filósofo alemão busca contrapor sua teoria à utopia do progresso ilimitado, representada pela obra de Ernst Bloch, "O Princípio Esperança", visto que a ideia de progresso sem limites constitui uma ameaça à própria existência humana e ao meio ambiente, uma vez que não leva em conta as possíveis consequências negativas das inovações tecnológicas (Fonseca, 2012).

O temor, segundo Jonas, surge não de uma visão pessimista da humanidade, mas da percepção de que a ação humana possui efeitos irreversíveis. Dessa forma, a responsabilidade deve preceder o agir, funcionando como um "saber de possibilidades" que permite a avaliação cuidadosa das consequências das ações humanas. Para ele, a responsabilidade deve ser entendida como um compromisso ético que transcende o interesse individual e se estende para as gerações futuras e para toda a vida no planeta (Fonseca, 2012).

A Ética da Responsabilidade de Jonas também encontra suporte no conceito de cuidado, que pode ser compreendido a partir da fábula-mito do "Cuidado", referida em "Ser e Tempo" de Heidegger. Esse mito revela a essência do ser humano como um ser de responsabilidade, cuja existência está intrinsecamente ligada à necessidade de zelar pelo outro e pelo mundo. Assim, Jonas propõe que a responsabilidade se torne a virtude central da ética contemporânea, substituindo a tradicional ética da reciprocidade (Fonseca, 2012).

Outro elemento central da teoria de Jonas é a assimetria da responsabilidade, influenciada pelo pensamento de Emmanuel Levinas. O exemplo paradigmático dessa assimetria é a relação entre pais e filhos: o recém-nascido é completamente vulnerável e incapaz de retribuir os cuidados que recebe, mas sua mera existência impõe aos pais a obrigação moral de protegê-lo. Jonas amplia essa ideia para incluir toda a vida no planeta, afirmando que a responsabilidade deve abarcar não apenas os seres humanos atuais, mas também as futuras gerações e os ecossistemas (Fonseca, 2012).

Portanto, verifica-se que sua proposta consiste em estabelecer uma moralidade que transcenda a relação imediata entre indivíduos e passe a

considerar o impacto das ações humanas no futuro da humanidade e do meio ambiente, redefinindo a ética como uma responsabilidade global baseada na prudência e na necessidade de proteger o futuro da vida na Terra.

## **4.2 A aplicação da Ética da Responsabilidade na regulação da aplicação da inteligência artificial no Direito**

Aplicando a perspectiva de Jonas às IAs no campo jurídico, percebe-se a necessidade de exigência de que o desenvolvimento e a implementação de sistemas de IA sejam acompanhados de um compromisso ético com a prevenção de injustiças e a mitigação de riscos sistêmicos.

Assim, no que concerne a ideia de regulamentação da IA no sistema jurídico, a ética de Jonas oferece um modelo de responsabilidade que pode orientar a criação de consensos regulatórios que priorizem a precaução, a sustentabilidade e a proteção dos direitos fundamentais, tanto para as gerações presentes quanto para as futuras (Fonseca, 2019).

Aplicando a perspectiva do aludido filósofo ao uso de IA no sistema jurídico, é assertivo afirmar que a Ética da Responsabilidade exige uma regulamentação que antecipe e evite possíveis danos que a IA possa causar à justiça, aos direitos humanos e à coesão social, implicando na criação de normas e diretrizes que não apenas lidem com os problemas presentes, mas que também considerem os impactos de longo prazo que a automação e a IA podem trazer (Fonseca, 2012).

Nessa toada, a Ética da Responsabilidade de Jonas sugere que, na criação de consensos regulatórios para a IA, deve-se adotar o princípio da precaução, especialmente em relação a tecnologias cujo impacto total ainda é desconhecido ou incerto (MARTSENKO, 2022). Por efeito, os reguladores devem errar pelo lado da cautela, garantindo que quaisquer inovações tecnológicas que possam afetar os direitos fundamentais sejam sujeitas a um exame rigoroso antes de serem amplamente implementadas. Isso poderia incluir requisitos mais rigorosos para a validação de algoritmos, auditorias de impacto ético e social, e testes extensivos para assegurar que as decisões automatizadas não comprometam a justiça e a equidade.

A falta de clareza na responsabilidade por decisões tomadas por

sistemas de IA é uma das maiores preocupações na filosofia do direito, especialmente porque a lógica de muitos algoritmos não é facilmente compreensível (MAYDANYK et al., 2021).

Dessa maneira, sob a ótica da ética de Jonas, a regulação da IA deve assegurar que todas as partes envolvidas na criação, desenvolvimento e aplicação da tecnologia sejam responsabilizadas não apenas pelos danos diretos, mas também pelos efeitos indiretos e de longo prazo que suas ações podem gerar. Isso pode incluir a imposição de padrões de segurança elevados, o monitoramento contínuo do impacto das tecnologias de IA, e a exigência de mecanismos de reparação justos para aqueles que possam ser prejudicados por decisões automatizadas (MIGUEL et al., 2020).

Ademais, a teoria de Jonas também implica a necessidade de um contratualismo ético, no qual a regulamentação da IA seja vista como um contrato social entre todos os membros da sociedade. Por efeito, exige-se que as políticas de IA sejam desenvolvidas de maneira inclusiva e democrática, envolvendo não apenas especialistas técnicos, mas também representantes da sociedade civil, grupos vulneráveis, operadores do direito e o público em geral (JONAS, 2006).

O objetivo é criar um consenso regulatório que reflita um compromisso compartilhado de proteger o bem-estar social e garantir que a inovação tecnológica seja usada de maneira que respeite a dignidade humana e os direitos fundamentais, evidenciando que a Ética da Responsabilidade incentiva um processo regulatório transparente e participativo, onde todas as partes interessadas tenham uma voz ativa e onde as normas sejam constantemente revisadas e adaptadas à luz das modificações sociais contemporâneas.

Por exemplo, a materialização dessa precaução na prática regulatória pode incluir requisitos mais rigorosos para a validação de algoritmos, auditorias de impacto ético e social, e testes extensivos para assegurar que as decisões automatizadas não comprometam a justiça e a equidade. Além disso, deve-se considerar a necessidade de avaliação contínua desses sistemas, dado que a IA opera em um ambiente dinâmico, no qual os modelos de aprendizado de máquina podem evoluir e, eventualmente, apresentar novas formas de viés ou comportamento inesperado.

A prestação de contas também exige a definição de um arcabouço

normativo que estabeleça direitos e garantias processuais para os indivíduos afetados por decisões automatizadas, assegurando que possam contestar e revisar tais decisões sempre que necessário. Para tanto, a regulamentação deve prever a obrigatoriedade de que as decisões judiciais baseadas em IA sejam explicáveis e justificáveis, permitindo que cidadãos e operadores do Direito compreendam os critérios utilizados na tomada de decisões algorítmicas.

Contudo, esse requisito fundamental de explicabilidade e previsibilidade entra em conflito direto com a natureza dos sistemas de IA baseados em aprendizado profundo, os quais, apesar de sua alta capacidade de análise e predição, apresentam o problema ainda não resolvido da opacidade algorítmica, impedindo a plena compreensão da lógica ou os fundamentos exatos que levaram a tais conclusões (Valença, 2007).

Nesse sentido, a regulamentação deve promover um modelo híbrido, no qual a IA atue como ferramenta auxiliar à decisão judicial, fornecendo análises e recomendações que possam ser interpretadas, questionadas e, se necessário, rejeitadas pelo julgador humano (Nunes; Marques, 2018).

Isso significa que os sistemas devem ser desenvolvidos com mecanismos que permitam a extração de justificativas inteligíveis para os resultados produzidos, viabilizando a realização de auditorias e inspeções regulares para garantir que o processo decisório automatizado respeite os preceitos constitucionais e os princípios fundamentais do Direito.

Por fim, a adoção de um modelo regulatório baseado na Ética da Responsabilidade impõe a necessidade de revisão periódica das normas aplicáveis à IA no Direito, assegurando que eventuais avanços tecnológicos permitam maior explicabilidade e previsibilidade conforme seu surgimento.

Dessa forma, garante-se que a evolução da inteligência artificial no campo jurídico ocorra dentro de um marco normativo sólido, que privilegie a transparência, a equidade e a proteção dos direitos fundamentais, afastando, assim, os riscos inerentes à utilização irrestrita da IA no campo do Direito.

## 5. CONSIDERAÇÕES FINAIS

Ao longo desta pesquisa, ficou evidente que o pensamento de Hans Jonas oferece uma contribuição significativa e singular para a discussão ética no contexto contemporâneo. Sua perspectiva ética da responsabilidade enfatiza não apenas as consequências imediatas das ações humanas, mas principalmente as implicações futuras e globais das inovações tecnológicas, revelando-se especialmente pertinente no contexto da IA, onde as decisões possuem efeitos que transcendem o presente e impactam o futuro de maneira significativa.

A Ética da Responsabilidade, conforme formulada por Jonas, defende a necessidade de adotar uma postura cautelosa frente à implementação e expansão das tecnologias, particularmente aquelas cuja complexidade pode gerar consequências imprevisíveis, como ocorre com a IA no campo jurídico. Sob essa ótica, constata-se que é imperativo introduzir mecanismos regulatórios preventivos e prudentes, capazes de antecipar os possíveis riscos e de assegurar que a IA seja utilizada em harmonia com os valores éticos fundamentais, preservando assim o bem-estar das gerações atuais e futuras.

Neste sentido, as conclusões deste estudo reforçam que uma regulamentação eficaz da IA aplicada ao Direito deve ser fundamentada em princípios que garantam a transparência dos algoritmos, a explicabilidade das decisões automatizadas, e a responsabilização clara e objetiva dos agentes envolvidos. Além disso, é essencial assegurar que tais ferramentas tecnológicas sejam vistas como complementos ao julgamento humano, jamais como substitutos plenos da decisão judicial.

Ademais, constatou-se que o princípio da responsabilidade, ao preconizar um compromisso ético proativo e preventivo, oferece uma sólida base dogmática capaz de orientar a construção de um arcabouço regulatório que contemple auditorias periódicas, mecanismos de supervisão contínua e a implementação de protocolos claros para a proteção de dados e privacidade dos cidadãos. Assim, promove-se uma utilização ética e segura da IA no âmbito jurídico, prevenindo violações aos direitos fundamentais e preservando a confiança pública nas instituições judiciais.

Portanto, em resposta ao problema central proposto nesta pesquisa, conclui-se que a perspectiva filosófica de Hans Jonas, com sua ênfase no princípio da responsabilidade e na precaução, constitui-se em uma base normativa adequada e eficaz para regulamentar a Inteligência Artificial no Direito. Por fim, sustenta-se que apenas com uma abordagem ético-filosófica robusta, como a proposta por Jonas, será possível enfrentar os desafios contemporâneos impostos pela IA, garantindo um desenvolvimento tecnológico que respeite os direitos fundamentais e os valores essenciais da justiça.

## REFERÊNCIAS

MAIA FILHO, Mamede Said; JUNQUILHO, Tainá Aguiar. Projeto Victor: perspectivas de aplicação da inteligência artificial ao direito. *Revista de Direitos e Garantias Fundamentais*, v. 19, n. 3, p. 218-237, 2018. Disponível em: <https://dialnet.unirioja.es/servlet/articulo?codigo=8697581>.

GÓRRIZ, Juan M. et al. Artificial intelligence within the interplay between natural and artificial computation: Advances in data science, trends and applications. *Neurocomputing*, v. 410, p. 237-270, 2020.

COUTINHO, Maria Gracielly F.; FERNANDES, Marcelo Augusto Costa. Proposta de implementação de rede neural artificial em hardware para aprendizagem profunda. In: Congresso Brasileiro de Automática-CBA. 2019. Disponível em: [https://www.sba.org.br/open\\_journal\\_systems/index.php/cba/article/view/738](https://www.sba.org.br/open_journal_systems/index.php/cba/article/view/738).

DE OLIVEIRA QUEIROZ, Júlia; GONÇALVES, Jonas Rodrigo; DA COSTA, Danilo. A questão digital: a inteligência artificial e o futuro da advocacia. *Revista JRG de Estudos Acadêmicos*, v. 7, n. 15, p. e151398-e151398, 2024. Disponível em: <https://revistajrg.com/index.php/jrg/article/view/1398>.

STEMLER, Ivan Sasha Viana. O processo criativo na era da inteligência artificial. 2023. Disponível em: <https://bdm.unb.br/handle/10483/37884>.

SILVA, Alexandre Pacheco da et al. Formando a advocacia do presente e do futuro: habilidades e perspectivas de atuação: destaques e tendências. *CEPI FGV Direito SP*, 2023. Disponível em: <https://repositorio.fgv.br/items/922b4466-d1ce-46a2-a5d2-d7139fc20756>.

ELIAS, Jonas da Silva et al. Recuperação e Análise de Jurisprudências como Suporte a Elaboração de Defesas em Processos Jurídicos por meio de Modelos de Linguagem. 2024. Disponível em: <https://repositorio.ufsc.br/handle/123456789/255787>.

CUEVA, Ricardo Villas Bôas. Inteligência artificial no Judiciário. In: WOLKART, Erik Navarro; NUNES, Dierle; LUCON, Paulo Henrique dos Santos (org.). *Inteligência artificial e direito processual*. Salvador: JusPodivm, 2021. Disponível em: <https://bdjur.stj.jus.br/jspui/handle/2011/140140>.

TAVARES-PEREIRA, Sebastião. Processo eletrônico, máxima automação, extraoperabilidade, imaginalização mínima e máximo apoio ao juiz: ciberprocesso. *Revista do Tribunal Regional do Trabalho da 13ª Região, João Pessoa*, v. 16, n. 1, 2009. Disponível em: <http://bdjur.stj.jus.br/dspace/handle/2011/35515>.

XU, Zichun. Human Judges in the Era of Artificial Intelligence: Challenges and Opportunities. In *Applied Artificial Intelligence*, v. 36, n. 1, 2022. Disponível em: <https://doi.org/10.1080/08839514.2021.2013652>.

RIBEIRO, João Luiz Vieira. Personalização da inteligência artificial: novo

paradigma jurídico. 2020. Disponível em:  
<https://repositorio.uniceub.br/jspui/handle/prefix/15192>.

PEIXOTO, Fabiano Hartmann; SILVA, Roberta Zumblick Martins da. Inteligência artificial e direito. Curitiba: Alteridade, 2019.

NASCIMENTO, Ivoney Severina de Melo Pereira do. Inteligência artificial e Direito: potencial da IA como ferramenta otimizadora da prestação jurisdicional voltada à superação das desigualdades sob a perspectiva do Estado de Coisas Inconstitucional. 2023. Disponível em:  
<http://www.rlbea.unb.br/jspui/handle/10482/46745>.

PAIVA, Mariana Filipa Menezes de. Inteligência artificial e o direito. 2022. Dissertação de Mestrado. Disponível em:  
<http://repositorio.ulusiada.pt/handle/11067/6803>.

TAKAKI, Lucas Mendonça. A incorporação de inteligência artificial no judiciário brasileiro: análise dos sistemas implementados na cadeia jurisdicional. 2022. Trabalho de Conclusão de Curso (Graduação em Direito) – Universidade XYZ, São Paulo, 2022. Disponível em:  
<https://repositorio.uniceub.br/jspui/handle/prefix/16391>.

DRAGALZEV, Gabriel Andreas. Inteligência artificial (IA): implicações acerca da sua utilização no âmbito jurídico. 2024. Disponível em:  
<https://repositorio.ueg.br/jspui/handle/riueg/4301>.

EHRHARDT JR, Marcos; SILVA, Gabriela Buarque Pereira. Contratos e Algoritmos: Alocação de Riscos Discriminação e Necessidade de Supervisão por Humanos. JURISMAT, n. 13, p. 33-33, 2021. Disponível em:  
<https://revistas.ulusofona.pt/index.php/jurismat/article/download/8153/4869>.

BASTOS, Camila Arruda Vidal; DA FONSÊCA, Mariana Guedes Duarte; VALENÇA, Manuela Abath. Morosidade, razoável duração do processo e acesso à justiça: Análise de tempo no sistema de justiça criminal em Recife. Revista Brasileira de Direitos Fundamentais & Justiça, v. 5, n. 17, p. 178-206, 2011. Disponível em: <https://dfj.emnuvens.com.br/dfj/article/view/349>.

JACOBSEN, Gilson; DE MACEDO DIAS, Bruno. Smart dispute resolution: Artificial intelligence to reduce litigation. Suprema-Revista de Estudos Constitucionais, v. 3, n. 1, p. 391-414, 2023. Disponível em:  
<https://suprema.stf.jus.br/index.php/suprema/article/view/215>

BAKSHI, Nitin; KIM, Jeunghyun; RANDHAWA, Ramandeep S. Service Operations for Justice-On-Time: A Data-Driven Queueing Approach. Available at SSRN 3764036, 2023. Disponível em:  
[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3764036](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3764036).

GONZÁLEZ, José Alfredo Godínez. La inteligencia artificial en el campo jurídico. Revista Académica CUNZAC, v. 6, n. 2, p. 96-103, 2023. Disponível em:  
<https://revistacunzac.com/index.php/revista/article/view/103>.

MAGALHÃES, Renato Vasconcelos. Automatización del raciocinio jurídico: perspectivas y límites en la aplicación de la inteligencia artificial al derecho. Burgos, 2004. 627 f. Tese (Doutorado) - Faculdade de Direito da Universidade de Burgos, p. 390-477. Disponível em: <https://investigacion.ubu.es/documentos/5db032a5299952180eb172d7>.

ALVES, Marco Antônio Sousa; ANDRADE, OM de. Da “caixa-preta” à “caixa de vidro”: o uso da explainable artificial intelligence (XAI) para reduzir a opacidade e enfrentar o enviesamento em modelos algorítmicos. *Direito Público*, v. 18, n. 100, 2022.

OLIANI, Louise Cristina Gonzaga; DE LIMA TEIXEIRA, Livia Solana Pfuetzenreiter. Moral e ética como critérios de decisões judiciais. *Revista Científica Sophia*, 2019. Disponível em:

NUNES, Dierle; MARQUES, Ana Luiza. Inteligência artificial e direito processual: vieses algorítmicos e os riscos de atribuição de função decisória às máquinas. *Revista de Processo*, v. 285, p. 421-447, novembro de 2018.

FERRARA, Emilio. Fairness and bias in artificial intelligence: A brief survey of sources, impacts, and mitigation strategies. *Sci*, v. 6, n. 1, p. 3, 2023. Disponível em: <https://www.mdpi.com/2413-4155/6/1/3>.

NERSESSIAN, David; MANCHA, Ruben. From automation to autonomy: legal and ethical responsibility gaps in artificial intelligence innovation. *Mich. Tech. L. Rev.*, v. 27, p. 55, 2020. Disponível em: <https://heinonline.org/HOL/LandingPage?handle=hein.journals/mttlr27&div=6&id=&page=>.

CARVALHO, Allan Pereira de. Viés algorítmico e discriminação: possíveis soluções regulatórias para o Brasil. 2020. Disponível em: <https://lume.ufrgs.br/handle/10183/221914>.

FERRAJOLI, Luigi. *Direito e razão: teoria do garantismo penal*. 3. ed. São Paulo: Editora Revista dos Tribunais, 2002.

ALVES, Jones Figueirêdo; PIMENTEL, Alexandre Freire Autor. Breves notas sobre os preconceitos decisoriais judiciais produzidos por redes neurais artificiais. *Revista da Faculdade de Direito da Universidade de Lisboa*, v. 63, p. 557-577, 2022. Disponível em: <https://repositorio.ul.pt/bitstream/10451/62129/1/Jones-Figueire%CC%82do-Alves-Alexandre-Freire-Pimentel.pdf>.

LEAL, Mônia Clarissa Hennig; PAULO, Lucas Moreschi. Algoritmos discriminatórios e jurisdição constitucional: os riscos jurídicos e sociais do impacto dos vieses nas plataformas de inteligência artificial de amplo acesso. *Revista de Direitos e Garantias Fundamentais*, v. 24, n. 3, p. 165-187, 2023. Disponível em: <https://sisbib.emnuvens.com.br/direitosegarantias/article/view/2311>.

GRGIĆ-HLAČA, N.; ZAFAR, M. B.; GUMMADI, K. P.; WELLER, A. Beyond Distributive Fairness in Algorithmic Decision Making: Feature Selection for Procedurally Fair Learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, [S. l.], v. 32, n. 1, 2018. DOI: 10.1609/aaai.v32i1.11296. Disponível em: <https://ojs.aaai.org/index.php/AAAI/article/view/11296>.

WARMLING, Natália Nicoski; BERNARDES, Marciele Berger; SANTOS, Paloma Maria. A lei de acesso à informação e o dever de transparência: uma análise do exercício da transparência passiva no âmbito dos Tribunais de Justiça brasileiros. *Eletrônico e Inclusão Digital*, p. 63, 2014.

VALENÇA, Mêuser Jorge Silva; LUDERMIR, Teresa Bernarda. Explicando a relação entre as variáveis de uma rede neural: Iluminando a “Caixa Preta”. XVII Simpósio Brasileiro de Recursos Hídricos, São Paulo, 2007. Disponível em: <[https://abrh.s3.sa-east-1.amazonaws.com/Sumarios/19/ea13ea2206546d31185a6ca6c7682074\\_a5d30c26854a6a7884d6635272be8882.pdf](https://abrh.s3.sa-east-1.amazonaws.com/Sumarios/19/ea13ea2206546d31185a6ca6c7682074_a5d30c26854a6a7884d6635272be8882.pdf)>.

DE ALBUQUERQUE MARANHÃO, Juliano Souza; JUNQUILHO, Tainá Aguiar; TASSO, Fernando Antonio. Transparência sobre o emprego de Inteligência Artificial no Judiciário: um modelo de governança. *Suprema-Revista de Estudos Constitucionais*, v. 3, n. 2, p. 145-187, 2023. Disponível em: <<https://legalgroundsinstitute.com/wp-content/uploads/2023/12/document-2.pdf>>.

GRIPON, Vincent; LASSANCE, Carlos; HACENE, Ghouthi Boukli. DecisiveNets: Training Deep Associative Memories to Solve Complex Machine Learning Problems. *arXiv preprint arXiv:2012.01509*, 2020. Disponível em: <<https://arxiv.org/abs/2012.01509>>.

FIORATTO, Débora Carvalho; DE CARVALHO DIAS, Ronaldo Brêtas. A conexão entre os princípios do contraditório e da fundamentação das decisões na construção do Estado Democrático de Direito. *Revista Eletrônica de Direito Processual*, v. 5, n. 5, 2010. Disponível em: <<https://www.e-publicacoes.uerj.br/redp/article/view/23088>>.

RUDIN, Cynthia; WANG, Caroline; COKER, Beau. The age of secrecy and unfairness in recidivism prediction. *Harvard Data Science Review*, v. 2, n. 1, p. 1, 2020. Disponível em: <<https://arxiv.org/abs/1811.00731>>.

ADITYA, P.; PAL, Mayukha. Local interpretable model agnostic shap explanations for machine learning models. *arXiv preprint arXiv:2210.04533*, 2022. Disponível em: <<https://arxiv.org/abs/2210.04533>>.

DEEKS, Ashley. The judicial demand for explainable artificial intelligence. *Columbia Law Review*, v. 119, n. 7, p. 1829-1850, 2019. Disponível em: <<https://www.jstor.org/stable/26810851>>.

BRKAN, M. Do Algorithms Rule the World? Algorithmic Decision-Making in the

Framework of the GDPR and Beyond. 2017. Disponível em: <<https://dx.doi.org/10.2139/SSRN.3124901>>.

FONSECA, Aline Klayse dos Santos et al. A prevenção na responsabilidade civil e a imputação pela formação do estado de danosidade. 2017. Disponível em: <[https://www.repositorio.ufpa.br/bitstream/2011/12131/1/Dissertacao\\_Prevencao\\_ResponsabilidadeCivil.pdf](https://www.repositorio.ufpa.br/bitstream/2011/12131/1/Dissertacao_Prevencao_ResponsabilidadeCivil.pdf)>.

MCDONALD, Louisa. AI Systems and Liability: An Assessment of the Applicability of Strict Liability & A Case for Limited Legal Personhood for AI. *St Andrews Law Journal*, v. 3, n. 1, p. 5-21, 2023. Disponível em: <<https://ojs.st-andrews.ac.uk/index.php/STALJ/article/view/2645>>.

REED, Chris. How should we regulate artificial intelligence?. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, v. 376, n. 2128, p. 20170360, 2018. Disponível em: <<https://royalsocietypublishing.org/doi/full/10.1098/rsta.2017.0360>>.

BENHAMOU, Yaniv; FERLAND, Justine. Artificial intelligence & damages: Assessing liability and calculating the damages. *Leading Legal Disruption: Artificial Intelligence and a Toolkit for Lawyers and the Law*, Forthcoming, 2020. Disponível em: <[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3535387](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3535387)>.

GARCÍA-VIGIL, José L. Reflections around ethics, human intelligence and artificial intelligence. *Gaceta medica de Mexico*, v. 157, n. 3, p. 298-301, 2021. Disponível em: <[https://web.archive.org/web/20220519080148id\\_/https://www.gacetamedicade-mexico.com/files/gmm\\_uk\\_21\\_157\\_3\\_298-301.pdf](https://web.archive.org/web/20220519080148id_/https://www.gacetamedicade-mexico.com/files/gmm_uk_21_157_3_298-301.pdf)>.

MARTIN, Kirsten. Designing ethical algorithms. *MIS Quarterly Executive* June, 2019. Disponível em: <[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3056692](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3056692)>.

BASHAYREH, Mohammad H.; TABBARA, Amer; SIBAI, Fadi N. The Need for a Legal Standard of Care in the AI Environment. *Sriwijaya Law Review*, v. 7, n. 1, p. 73-86, 2023. Disponível em: <<http://journal.fh.unsri.ac.id/index.php/sriwijayalawreview/article/view/1507>>.

CATH, Corinne. Governing artificial intelligence: ethical, legal and technical opportunities and challenges. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, v. 376, n. 2133, p. 20180080, 2018. Disponível em: <<https://royalsocietypublishing.org/doi/full/10.1098/rsta.2018.0080>>.

PAUL, Regine. European artificial intelligence “trusted throughout the world”: Risk-based regulation and the fashioning of a competitive common AI market. *Regulation & Governance*, 2023. Disponível em: <<https://onlinelibrary.wiley.com/doi/full/10.1111/rego.12563>>.

TEUBNER, Gunther. Digital personhood? The status of autonomous software agents in private law. *The Status of Autonomous Software Agents in Private Law* (May 11, 2018), 2018. Disponível em: <[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3177096](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3177096)>.

IANNÌ, A.; MONTEROSSÌ, M. W. Artificial autonomous agents and the question of electronic personhood: a path between subjectivity and liability. 2017. Disponível em: <https://dx.doi.org/10.1080/10383441.2017.1558611>

CURZON, J.; KOSA, T. A.; AKALU, R.; EL-KHATIB, K. Privacy and Artificial Intelligence. 2021.

BIELOVA, M.; BYELOV, D. Challenges and threats of personal data protection in working with artificial intelligence. 2023.

MARTIN, K.; RUBINSTEIN, I. S. Data Privacy and the Ethical Use of Artificial Intelligence. 2020.

CAVOUKIAN, A. Privacy by Design: The 7 Foundational Principles. 2019.

DONEDA, D.; MENDES, L. S.; SOUZA, C.; GOMES DE AN, N. N. M. B. Considerações iniciais sobre inteligência artificial, ética e autonomia pessoal. 2018.

GASPARIAN, D. E.; TURKO, D. S.; BESSCHETNOVA, E. V. Ethical Possibilities of Using Artificial Intelligence: Cultural and Philosophical Analysis. 2023. Disponível em: <https://dx.doi.org/10.21146/0042-8744-2023-9-93-102>.

ROGERS, J.; BELL, F. The Ethical AI Lawyer: What is Required of Lawyers When They Use Automated Systems?. 2019.

CHISTILINA, D. O. Using the Capabilities of Artificial Intellect in Criminal Proceedings. 2021.

ZUCKERMAN, A. Artificial Intelligence in the Administration of Justice. 2020.

BEBIANO, F. N. Aplicação da Inteligência Artificial nos Conflitos Submetidos à Justiça Restaurativa: (Im)Possibilidade. 2022.

ENGELMANN, A. Algorithmic transparency as a fundamental right in the democratic rule of law. 2023.

RODRÍGUEZ, N. D.; SER, J.; COECKELBERGH, M.; LÓPEZ DE PRADO, M.; HERRERA-VIEDMA, E.; HERRERA, F. Connecting the Dots in Trustworthy Artificial Intelligence: From AI Principles, Ethics, and Key Requirements to Responsible AI Systems and Regulation. 2023.

ZHAO, H. Judicial Trials in the Context of Artificial Intelligence. 2021. Disponível em: <https://dx.doi.org/10.1088/1742-6596/1744/4/042175>.

FONSÊCA, F. Ética da responsabilidade em Hans Jonas e limites das éticas tradicionais (antropocêntricas). *Revista Eclesiástica Brasileira*, 2019. Disponível em: <https://dx.doi.org/10.29386/REB.V68I269.1467>.

GONÇALVES, I.; TORRES, E. K. P. Ética, tecnologia e responsabilidade: uma abordagem a partir do pensamento de Hans Jonas. *Cadernos do PET-FIL*, 2020. Disponível em: <https://dx.doi.org/10.26694/CADPETFIL.V10I20.11407>.

LEAL, A.; DURANTE, D. O princípio responsabilidade em Hans Jonas como proposta de ética para uma sociedade sustentável. *Revista de Direito Ambiental*, 2021. Disponível em: <https://dx.doi.org/10.5380/dma.v58i0.71625>.

YARA, O.; BRAZHEYEV, A.; GOLOVKO, L.; BASHKATOVA, V. Legal Regulation of the Use of Artificial Intelligence: Problems and Development Prospects. 2021. Disponível em: <https://dx.doi.org/10.14207/EJSD.2021.V10N1P281>.

MARTSENKO, N. Artificial Intelligence and Human Rights: a Scientific Review of Impacts and Interactions. 2022. Disponível em: <https://dx.doi.org/10.31648/sp.8245>.

RODRÍGUEZ, Antonio Luis Terrones. *Inteligencia artificial y ética de la responsabilidad*. 2018.

JONAS, Hans. *O princípio responsabilidade: ensaio de uma ética para uma civilização tecnológica*. Rio de Janeiro: PUC Rio, 2006.

BASTOS, Camila Arruda Vidal; DA FONSÊCA, Mariana Guedes Duarte; VALENÇA, Manuela Abath. Morosidade, razoável duração do processo e acesso à justiça: Análise de tempo no sistema de justiça criminal em Recife. *Revista Brasileira de Direitos Fundamentais & Justiça*, v. 5, n. 17, p. 178-206, 2011.

RUSSEL, Stuart; NORVIG, Peter. *Inteligência Artificial*. 2. Ed. Rio de Janeiro: Campos, 2004.

SIQUEIRA, Tiago de Oliveira. *Inteligência artificial e sua influência no judiciário brasileiro*. In: *Educação e Tecnologias em Debate: Perspectiva sob diferentes áreas do conhecimento*. São Carlos: Pedro & João Editores, 2020. p. 87-94.

BRENE, Bárbara Cheder. *A UTILIZAÇÃO DE INTELIGÊNCIA ARTIFICIAL NO PODER JUDICIÁRIO COMO ALTERNATIVA AO PROBLEMA DA MOROSIDADE PROCESSUAL*. 2021.

FERREIRA FILHO, Moacir Ferreira Filho Moacir. *A DESUMANIZAÇÃO DO DIREITO: UMA ABORDAGEM ACERCA DO USO DA TECNOLOGIA NO AMBIENTE JURÍDICO*. *Revista Interfaces*, v. 15, n. 11, 2023.

DE MARCO, Cristhian Magnus; LEMES, Mariana Carolina; DE PAULA CHIESSE, Daniel Roxo. *ALGORITMOS: códigos invisíveis (d) e injustiça*, 2020.

RODRÍGUEZ, Antonio Luis Terrones. *Ética para la inteligencia artificial*

sostenible. Arbor, v. 198, n. 806, p. a683-a683, 2022.

NODARI, Paulo César. Ética da responsabilidade em Hans Jonas. Minas Gerais, Belo Horizonte, 2014.

FONSECA, Flaviano Oliveira. Hans Jonas: ética para a civilização tecnológica. Cadernos de Ciências Sociais Aplicadas, 2012.

BARBOSA, Isaias Mendes. A ética da responsabilidade em Hans Jonas: crítica à modernidade e novos caminhos de atuação. Annales Faje, v. 2, n. 2, p. 161-168, 2017.

BARBOSA, Lucia Martins; PORTES, Luiza Alves Ferreira. A inteligência artificial. Revista Tecnologia Educacional [on line], Rio de Janeiro, n. 236, p. 16-27, 2023.