A RESPONSABILIDADE CIVIL PELOS DANOS CAUSADOS POR SISTEMAS DE INTELIGÊNCIA ARTIFICIAL: DESAFIOS ÉTICOS E JURÍDICOS NO DIREITO CONTEMPORÂNEO

CIVIL LIABILITY FOR DAMAGES CAUSED BY ARTIFICIAL INTELLIGENCE SYSTEMS: ETHICAL AND LEGAL CHALLENGES IN CONTEMPORARY LAW

Alani Leticia Torres dos Santos[1]
Beatriz Aparecida Merenda Santos[2]

RESUMO

O desenvolvimento da Inteligência Artificial (IA) tem conferido a sistemas de computador a capacidade de aprender e atuar de forma autônoma, independentemente de instruções específicas humanas. Tal evolução tecnológica é marcada por sofisticação técnica e escala, tornando-se um desafio aos quadros jurídicos clássicos, sobretudo no que se refere à responsabilidade civil. A IA opera com base na inteligência sensível, que se restringe ao processamento de dados e reconhecimento de padrões, sendo incapaz de apreender valores, consciência moral ou verdade, características exclusivas da inteligência senciente humana. Essa distinção epistemológica torna problemática a aplicação de regimes de responsabilidade subjetiva, que exigem a comprovação de dolo ou culpa, uma vez que a máquina não possui intencionalidade ou consciência moral. A imprevisibilidade e a autoaprendizagem da IA, que podem levar a danos não antecipados por seus criadores (o chamado *black-box*), dificultam o estabelecimento do nexo de causalidade e do ônus da prova de defeito. A realidade da IA reforça a necessidade de revisão dos pressupostos clássicos e aponta para regimes de responsabilidade objetiva fundados em esferas de risco ou modelos econômicos de causalidade. A União Europeia, através da Resolução do Parlamento Europeu de 16 de fevereiro de 2017, sugere a adoção de um sistema de responsabilidade objetiva para quem está mais bem posicionado para minimizar riscos e oferecer garantias (*deep-pocket*). As soluções propostas incluem a criação de seguros obrigatórios, complementados por fundos de compensação. No Brasil, a regulação em discussão (e.g., PL 2338/2023 e LGPD) enfatiza a importância da supervisão humana (*human-in-the-loop*), da transparência e da *accountability* como mecanismos para manter o controle ético sobre os sistemas de alto risco. Conclui-se que o direito deve evoluir, adotando abordagens complementares que priorizem a reparação do dano e reafirmem a centralidade da inteligência senciente humana sobre a técnica.

Palavras-chave: Responsabilidade Civil. Inteligência Artificial. Robótica.

Supervisão Humana. Risco.

ABSTRACT

The development of Artificial Intelligence (AI) has given computer systems the ability to learn and act autonomously, independently of specific human instructions. This technological evolution, marked by technical sophistication and scale, challenges classical legal frameworks, especially civil liability. AI operates based on sensible intelligence, which

[1] Atua na área trabalhista na Vedana Sociedade de Advogados (desde 08/2025). Atuou no escritório Alloisio Neto Advogados Associados (03/2025 a 08/2025) e no Tribunal de Justiça de São Paulo, 5ª Vara Previdenciária (11/2023 a 02/2025). Realizou iniciação científica (2024) sobre a exposição e a monetização infantil nas redes sociais.

[2] Atua atualmente no escritório Marchiori, Sachet, Barros e Dias, em processos tributários. Atuou no Tribunal Regional Federal da 3ª Região (02/2024 a 04/2025). Realizou iniciação científica (2024) sobre responsabilidade civil e inteligência artificial.

is limited to data processing and pattern recognition, incapable of apprehending values, moral consciousness, or truth, exclusive characteristics of human sentient intelligence. This epistemological distinction makes the application of subjective liability regimes problematic, requiring proof of intent or fault, since the machine lacks intentionality. The unpredictability and self-learning capacity of AI, which can lead to damages unforeseen by its creators (the so-called *black-box*), complicate establishing the causal link and the burden of proof of defect. The reality of AI reinforces the need to review classical premises and points towards objective liability regimes based on risk spheres or economic causality models. The European Union, through the European Parliament Resolution of February 16, 2017, suggests adopting an objective liability system for those best positioned to minimize risks and offer guarantees (*deep-pocket*). Proposed solutions include mandatory insurance, complemented by compensation funds. In Brazil, the regulation under discussion (e.g., PL 2338/2023 and LGPD) emphasizes the importance of human supervision (*human-in-the-loop*), transparency, and accountability as mechanisms to maintain ethical control over high-risk systems. It is concluded that the law must evolve, adopting complementary approaches that prioritize damage repair and reaffirm the centrality of human sentient intelligence over technology.

Key words: Civil Liability. Artificial Intelligence. Robotics. Human Supervision.

Risk.

1. INTRODUÇÃO

A emersão da Inteligência Artificial (IA) representa uma transformação disruptivas da atualidade, afetando diversos setores como saúde, finanças e segurança. A capacidade da IA de transformar características da inteligência humana em processos automatizados, executando tarefas de forma autônoma, aprendendo com dados e tomando decisões, redefine totalmente a relação da humanidade com a tecnologia. Contudo, tal sofisticação técnica e a grande escala com que a IA opera tornam incontornável a análise

das múltiplas dimensões em que desafia os quadros jurídicos tradicionais, especialmente a responsabilidade civil.

Dito isso, um dos principais desafios é a definição de um responsável pelos danos causados por esses sistemas. As normas jurídicas tradicionais foram elaboradas fazendo relação com tecnologias precedentes, ou seja, estruturaram-se sobre a responsabilidade de indivíduos ou organizações por ações humanas. Na atual realidade, tomada por veículos autônomos, agentes de software, robôs de assistência e a Internet das coisas, tornou-se necessária uma revisão dos pressupostos clássicos da responsabilidade civil.

O objetivo desta pesquisa é analisar as limitações dos regimes de responsabilidade civil, tanto subjetiva quanto objetiva, no contexto da IA, explorando como a doutrina e as propostas regulatórias internacionais, em especial a Resolução do Parlamento Europeu de 2017, buscam adaptar-se aos danos derivados dos atos independentes da IA. Além disso, o estudo visa destacar a importância da fundamentação filosófica e epistemológica da IA para

a construção de um marco jurídico adequado.

2. DESAFIOS EPISTEMOLÓGICOS E JURÍDICOS DA IA

A Inteligência Artificial (IA) é definida como uma classe de tecnologias computacionais que buscam replicar ou substituir o conhecimento e as formas de conhecer próprias dos seres humanos. Por lidar com noções como *inteligência*, *pensamento* e *conhecimento*, a IA se conecta diretamente com a filosofia, especialmente com a epistemologia. O principal desafio filosófico, portanto, reside no paradigma de conhecimento que sustenta a IA.

2.1 Inteligência Sensível vs. Inteligência Senciente

As concepções que fundamentam a IA derivam de uma compreensão empirista e mecanicista do conhecimento humano, que remonta à tradição inglesa e culmina na filosofia de David Hume (1711–1776). Tal tradição sustenta a teoria da inteligência sensível (*sensible intelligence*), segundo a qual o conhecimento tem origem nos sentidos: o cérebro (ou, por analogia, o processador da máquina) recebe informações do ambiente, processa-as e forma uma "imagem" representacional do mundo exterior.

Diante disso, o conhecimento que resulta de "impressões", que dão origem a "ideias" não passam cópias pálidas e sem vida das impressões diretas (Hume, 1748, sec.

12, pt. 3). A partir da lógica por traz da IA, as "impressões" correspondem aos dados brutos coletados pelos sensores, enquanto as "ideias" se traduzem nas estruturas de dados armazenadas na memória, que o algoritmo manipula para produzir respostas ou ações. Assim, a IA reproduz o esquema humano de separação entre sentir e conhecer, estruturando-se sobre um modelo modular típico da engenharia: sensores que captam o ambiente e processadores que interpretam e executam comandos.

Outro elemento central é o nominalismo, ou seja, a crença de que entidades abstratas não possuem existência real, sendo apenas nomes atribuídos a coleções de indivíduos concretos. Essa suposição, existente nos sistemas de IA generativa e *Large Language Models* (LLM), implicam que a IA não compreende conceitos universais ou valores abstratos, mas apenas padrões observáveis. Para o algoritmo, "justiça", "verdade" ou "direitos humanos" são apenas palavras correlacionadas estatisticamente, não realidades inteligíveis.

Finalmente, a IA opera segundo uma noção de causalidade como conjunção constante, também herdada de Hume: as relações causais são reduzidas a repetições observadas e quantificáveis, em analogia com o modo como algoritmos aprendem por meio

de *feedback loops* e inferências estatísticas (Hume, 1951, I.I.I, p. 3). A IA, portanto, permanece restrita à causalidade material e eficiente, sem alcançar as dimensões formal e final, essenciais à compreensão humana de sentido e propósito.

Em síntese, a IA se desprende sobre um modelo epistemológico que confunde processamento de informação com ato de conhecer, assumindo que a "inteligência sensível, presente nas maquinas, possui as mesmas capacidades da inteligência sentiente humana. Tal base filosófica não compreende suas limitações mais profundas: a incapacidade de perceber a realidade e a ausência de um conceito de verdade.

2.2 O Paradigma Humano: a Inteligência Sentiente

Em contraste com o modelo empírico-mecanicista, o ser humano conhece a partir de uma inteligência sentiente (*sentient intelligence*), conceito desenvolvido por Xavier Zubiri e retomado por Fowler. Nessa visão, a percepção e o conhecimento não são atividades separadas, mas um processo unificado, no qual o sujeito está em contato direto com a realidade (Zubiri, 1981, p. 406). O ser humano não apenas registra informações,

8 ZUBIRI, XAVIER. GOD AND DEIFICATION IN PAULINE THEOLOGY. LANHAM, MD: UNIVERSITY PRESS OF AMERICA, 1981.

7 FOWLER, Thomas. Computers, Computation, and the Limits of Computability. Washington, DC: XZF Technical Publications, 2021.

mas experimenta o real, interpretando-o e atribuindo-lhe valor, finalidade e sentido (Fowler, 2021, pp. 109-199).

É precisamente essa capacidade sentiente, de apreender o verdadeiro, o justo e o bom, que distingue a inteligência humana de qualquer forma de simulação algorítmica. A IA pode reconhecer padrões, mas não compreender significados; pode responder coerentemente, mas não discernir a verdade; pode replicar estilos criativos, mas não criar no sentido humano do termo.

4.3. INABILIDADE PARA A VERDADE E A CRIATIVIDADE

A partir dessa base filosófica limitada, emergem as falhas mais críticas da IA, que comprometem sua confiabilidade em campos como o Direito, a ética e a ciência. Os sistemas generativos frequentemente produzem informações falsas, esse fenômeno é conhecido como "alucinação" e reproduz vieses ideológicos decorrentes dos dados e das intenções de seus programadores. Casos recentes, como o do Google Gemini, demonstram que tais sistemas estão longe de ser neutros, refletindo preconceitos e interesses embutidos

em seus conjuntos de treinamento.

Essas falhas decorrem de uma operação puramente estatística, voltada à persuasão e verossimilhança, e não à verdade. Como observa Fowler, a IA é construída para ser convincente, não verdadeira: baseia-se em probabilidades linguísticas, não em juízos epistemológicos. Assim, a máquina não conhece o conteúdo que produz, apenas o reproduz de modo plausível. Sua "competência discursiva" não se converte em competência cognitiva, razão pela qual, embora imite com precisão a linguagem humana, permanece ontologicamente incapaz de compreender significado ou veracidade.

A mesma limitação explica sua incapacidade criativa. A inteligência humana é radicalmente inovadora, capaz de formular teorias e transcender padrões; já a IA é prisioneira do passado, dependente de dados preexistentes e repetições estatísticas. Sua "originalidade" é apenas recombinação. Em última instância, tanto o viés quanto a ausência de criatividade resultam da impossibilidade de a IA operar fora do domínio dos dados e das correlações formais. Diferentemente da mente humana, que interpreta e projeta o novo, a inteligência artificial permanece confinada ao já dito: é uma razão sem consciência e sem verdade, incapaz de substituir a inteligência sentiente humana nas esferas da moral, da arte, da filosofia e do Direito.

3. O COLAPSO DOS PRESSUPOSTOS CLÁSSICOS DA RESPONSABILIDADE CIVIL

A crescente autonomia da IA impõe a revisão dos pressupostos clássicos da responsabilidade civil**.** O nexo de causalidade é o primeiro a ser desafiado: o anonimato, a complexidade e a imprevisibilidade dos sistemas autônomos dificultam a identificação do responsável e tornam o ônus da prova extremamente pesado para o lesado, sobretudo quanto à demonstração do defeito do produto.

A culpa também enfrenta impasses. No Direito português, o dever de indenizar baseia-se na culpa, salvo exceções legais. Contudo, atribuir culpa à IA é inviável, pois ela carece de intenção e consciência, elementos essenciais para o dolo ou a culpa. Assim, o dano causado por um robô não decorre de um processo volitivo próprio, mas de uma programação ou operação humana.

Dessa forma, a imputabilidade subjetiva só pode ser reconhecida na pessoa física ou jurídica vinculada ao robô. Como observa Antunes (2024), o autor do dano deve ser capaz de entender e querer no momento do ato lesivo. Portanto, a responsabilidade, no estado atual, permanece humana, ainda que mediada por sistemas autônomos.

5.1. Responsabilidade nas criações de IA

O principal fator que influencia a percepção de autoria, criação e responsabilidade do autor humano é o grau de assistência recebido, isto é, o nível de intervenção da IA ou de outro assistente no processo criativo. Quanto maior a intervenção da IA, menor tende a ser a percepção de autoria e responsabilidade humana; e quanto menor o auxílio, mais o autor é reconhecido como criador e responsável.

Essa relação inversa gera uma transferência perceptiva de autoria e culpa, em que a autoria simbólica, ligada à criatividade, intencionalidade e controle, se desloca do humano para a máquina conforme aumenta o grau de automatização.

Sob o ponto de vista jurídico, o grau de assistência é elemento essencial na relação entre o agente humano e a tecnologia. Não basta constatar o uso da IA: é preciso analisar o nível de influência da máquina sobre o processo decisório, a forma e o conteúdo do resultado. Tal análise é crucial tanto para definir a autoria (em propriedade intelectual) quanto para delimitar a responsabilidade civil.

Quando a IA atua com autonomia significativa, a presunção de autoria humana se enfraquece, pois, o criador perde o controle efetivo sobre o resultado. Nesses casos, falar em responsabilidade subjetiva do autor pode ser artificial ou injusto, já que se esvaziam os elementos clássicos da culpa.

Assim, o grau de assistência torna-se um novo parâmetro jurídico e ético, essencial para reconstruir os conceitos de autoria e responsabilidade na era da inteligência assistida. Ele permite distinguir o apoio instrumental (baixo grau de assistência) da coautoria funcional da IA (alto grau de assistência), exigindo revisão dos regimes de imputação e do dever de diligência.

Juridicamente, prevalece a visão de que a IA é uma ferramenta, não um sujeito de direito. A doutrina majoritária rejeita sua personalidade jurídica, afirmando que apenas pessoas físicas ou jurídicas humanas podem ser responsabilizadas. A IA, portanto, permanece no plano dos meios de execução, mesmo que dotada de autonomia funcional.

Contudo, essa visão enfrenta tensões crescentes. Em contextos de alta automação e opacidade algorítmica, a IA exerce funções decisórias e criativas que ultrapassam o controle técnico humano, aproximando-se de um agente funcional autônomo.

Surge, então, a questão central: se autoria e responsabilidade dependem do grau de contribuição causal e criativa, e não da natureza jurídica do agente, o Direito precisará repensar seus critérios de imputação. O princípio do primado da ferramenta, embora útil, pode tornar-se insuficiente diante de sistemas que já atuam como coprodutores de conteúdo e decisões**.**

5.2. Crítica ao Princípio da Limitação da Finalidade

O princípio da limitação da finalidade (*purpose limitation*), consagrado no Regulamento Geral de Proteção de Dados (GDPR) e em legislações correlatas, busca assegurar que os dados pessoais sejam processados apenas para fins específicos, legítimos e transparentes. No entanto, na era da IA, esse princípio se mostra cada vez mais insuficiente.

A natureza preditiva e adaptativa dos sistemas de IA extraem correlações ocultas e redefinem seus próprios objetivos com base em *machine learning*, desafiando a noção clássica de finalidade estável. A opacidade algorítmica impede que se determine com

precisão o propósito efetivo de cada processamento. Assim, o uso de dados extrapola o consentimento original, diluindo a fronteira entre finalidade legítima e exploração indevida.

O Direito, diante disso, enfrenta o dilema de conciliar inovação tecnológica e proteção ética, revisitando o princípio da limitação da finalidade à luz da complexidade informacional contemporânea. É necessário estabelecer novos marcos de responsabilidade e transparência algorítmica, que garantam a supervisão humana, a aplicabilidade dos resultados e o controle social sobre as decisões automatizadas.

A IA, enraizada na tradição da inteligência sensível, permanece limitada a uma epistemologia da representação e da repetição. Já o ser humano, guiado pela inteligência sentiente, conhece por contato com o real, pela atribuição de sentido e pela busca da verdade.

No campo jurídico, reconhecer essa diferença não é mera especulação filosófica, mas uma exigência prática: sem percepção da realidade e sem consciência moral, a IA não pode exercer julgamento justo. O desafio contemporâneo consiste, portanto, em desenvolver tecnologias que sirvam de suporte ao humano, construindo um Direito que continue a afirmar a centralidade da pessoa e da verdade em meio à lógica algorítmica.

5.3. A Incompatibilidade entre Opacidade (Black-Box) e a Responsabilidade Profissional

Para além das implicações sobre a finalidade do tratamento de dados, a opacidade algorítmica (*black-box problem*) gera impactos diretos sobre a responsabilidade profissional e a *accountability* (prestação de contas) dos agentes humanos. Profissionais clínicos possuem obrigações regulatórias de justificar suas decisões, mas um sistema opaco impede o usuário (clínico ou operador) de compreender e explicar como a recomendação da IA foi gerada.

Se o clínico não consegue compreender o raciocínio do sistema, ele não pode justificar plenamente suas ações. Essa incompatibilidade entre a opacidade da IA e o dever humano de prestação de contas leva à conclusão de que não é admissível, sob o prisma ético e jurídico, que um profissional utilize um Sistema de Inteligência Artificial (AIS) opaco na tomada de decisão.

A situação torna-se ainda mais complexa porque, ao redefinir seus próprios objetivos com base em *machine learning*, a IA reforça o caráter imprevisível e não auditável de seus resultados. Assim, o vínculo causal entre a ação humana e o dano se dilui, dificultando a imputação de culpa e comprometendo a própria noção de responsabilidade civil.

Diante da fragilidade do princípio da limitação da finalidade no contexto da reutilização de modelos, parte da doutrina propõe a criação de um "regime de *accountability* para modelos". Tal regime imporia obrigações de registro e rastreabilidade dos dados de treinamento, além da prova da finalidade original do modelo, de modo a permitir a reconstrução dos processos decisórios e a responsabilização efetiva dos agentes envolvidos. Esse controle ético permanente fortaleceria tanto a transparência quanto a confiança social no uso da IA.

Adicionalmente, deve-se considerar o risco sistêmico da "privacidade inferencial ou preditiva", no qual até dados anonimizados podem ser utilizados para inferir informações pessoais sobre terceiros. Esse fenômeno expande o dano para além do indivíduo, alcançando categorias inteiras de pessoas e ameaçando o núcleo do princípio da autodeterminação informacional.

Em síntese, a incompatibilidade entre a opacidade algorítmica e a responsabilidade profissional evidencia uma crise estrutural da *accountability* na era da IA. A ausência de explicabilidade técnica compromete o dever de diligência e a capacidade de justificação dos profissionais, exigindo, portanto, novos marcos regulatórios que imponham transparência, rastreabilidade e controle humano contínuo sobre os sistemas automatizados.

4. RISCOS ÉTICOS E REGULATÓRIOS DA INTELIGÊNCIA ARTIFICIAL: DO GDPR À REALIDADE BRASILEIRA

4.1 Fragilidade da Limitação de Finalidade e a Lacuna da Reutilização de Modelos

O princípio da limitação da finalidade, central no Regulamento Geral de Proteção de Dados da União Europeia (GDPR), busca assegurar que dados pessoais sejam coletados

e processados apenas para fins específicos e legítimos. No entanto, tal princípio

11 UNIÃO EUROPEIA. Resolução do Parlamento Europeu, de 16 de fevereiro de 2017, com recomendações à Comissão de Direito Civil sobre Robótica (2015/2103(I

mostra-se ineficaz diante da escala e complexidade da Inteligência Artificial. Após a anonimização dos dados de treinamento, a rastreabilidade da finalidade original se perde, rompendo o vínculo jurídico entre dado e propósito e inviabilizando o controle sobre usos subsequentes.

No Brasil, a Lei Geral de Proteção de Dados (Lei nº 13.709/2018 — LGPD) adota princípio idêntico em seu art. 6º, I, mas enfrenta o mesmo dilema. A anonimização (arts. 5º, XI, e 12) retira o dado do escopo de aplicação da lei, fragilizando a responsabilização em casos de reutilização indevida de conjuntos de dados anonimizados em modelos de IA. Assim como no GDPR, cria-se uma brecha pela qual modelos treinados com dados anonimizados podem ser reaplicados em contextos diversos, inclusive para fins discriminatórios, comerciais ou de vigilância, sem consentimento ou supervisão.

Um exemplo prático no contexto brasileiro ilustra a gravidade dessa fragilidade: bases de dados de saúde ou biometria coletadas sob pretexto de interesse público (como em programas de vacinação ou segurança) podem ser posteriormente reutilizadas por entes privados, como planos de saúde ou sistemas automatizados de recrutamento, sem transparência e sem controle efetivo pela Autoridade Nacional de Proteção de Dados (ANPD).

A lacuna regulatória não é exclusiva do Brasil. A própria União Europeia reconhece que tanto o GDPR quanto o AI Act tratam apenas dos riscos de uso da IA, mas não da reutilização secundária de modelos. Modelos treinados sob propósitos legítimos podem circular livremente e ser reaplicados em contextos distintos, gerando riscos sistêmicos e opacidade sobre as cadeias de responsabilidade. O PL 2338/2023 (Marco Legal da Inteligência Artificial) repete o mesmo vácuo: concentra-se no risco da aplicação da IA, mas não no risco estrutural da circulação de modelos e bases de treinamento.

Diante disso, parte da doutrina brasileira propõe a criação de um "regime de accountability para modelos", destinado a reconstituir a rastreabilidade e a finalidade jurídica dos sistemas de IA. Esse regime incluiria:

- dever de registro e rastreabilidade dos conjuntos de dados utilizados no treinamento;
- obrigação de auditoria independente e verificação periódica de conformidade;

14 UNIÃO EUROPEIA. RESOLUÇÃO DO PARLAMENTO EUROPEU, DE 16 DE FEVEREIRO DE 2017.

13 BRASIL. PROJETO DE LEI Nº 2338/2023 – MARCO LEGAL DA INTELIGÊNCIA ARTIFICIAL.

12 BRASIL. LEI Nº 13.709, DE 14 DE AGOSTO DE 2018. LEI GERAL DE PROTEÇÃO DE DADOS PESSOAIS (LGPD).

- exigência de prova da finalidade original, análoga ao que Fowler e outros teóricos europeus denominam de *"limitação da finalidade para modelos"*.

Tal estrutura busca restabelecer o vínculo perdido entre dado, modelo e propósito, evitando que a anonimização sirva de escudo para práticas de reuso opacas e juridicamente irresponsáveis.

6.2. Assimetria de poder informacional e riscos sistêmicos

Os sistemas de IA baseados em *machine learning* geram uma nova assimetria de poder informacional entre os detentores dos modelos e a sociedade. O risco é de natureza sistêmica, uma vez que mesmo dados anonimizados podem ser utilizados para inferir informações pessoais - fenômeno denominado pela doutrina de privacidade inferencial ou preditiva**.**

Dessa forma, qualquer indivíduo, ainda que fora da base de dados original, pode ser afetado por inferências estatísticas produzidas por modelos reutilizados. No Brasil, a Lei Geral de Proteção de Dados (LGPD) permanece centrada na tutela de direitos individuais, o que se mostra insuficiente para enfrentar riscos de natureza coletiva e difusa.

A assimetria de poder informacional entre grandes plataformas, como Google, Meta e OpenAI, e os cidadãos é crescente. A Autoridade Nacional de Proteção de Dados (ANPD), em suas diretrizes de 2024 sobre IA e proteção de dados, já reconheceu a necessidade de uma abordagem coletiva e setorial, inspirada em regimes de responsabilidade ampliada.

Ademais, embora o art. 21 da LGPD e o art. 20 (que trata da revisão humana de decisões automatizadas) constituam instrumentos relevantes, ainda se revelam insuficientes diante da opacidade dos modelos generativos e preditivos, os quais ultrapassam o escopo da decisão automatizada individual e passam a impactar grupos e categorias inteiras de pessoas.

6.3. Violação de privacidade e discriminação coletiva

16 BRASIL. LEI Nº 13.709, DE 14 DE AGOSTO DE 2018. LEI GERAL DE PROTEÇÃO DE DADOS PESSOAIS (LGPD).

15 FOWLER, Thomas. Computers, Computation, and the Limits of Computability. Washington, DC: XZF Technical Publications, 2021.

O risco ético e jurídico da IA não se limita à re-identificação de indivíduos, mas inclui discriminação e predição coletiva. Um modelo treinado com dados médicos, por exemplo, pode ser reutilizado para avaliar perfis de risco em seguros de vida ou processos seletivos, violando a finalidade original e os princípios da não discriminação.

No contexto brasileiro, o art. 6º, IX, da LGPD e o art. 5º, XLI, da Constituição Federal proíbem discriminação baseada em dados pessoais, saúde, gênero ou origem. No entanto, a prova da discriminação algorítmica ainda é um desafio. Faltam mecanismos processuais e técnicos para detectar e coibir o uso indireto de modelos enviesados, especialmente quando esses modelos foram treinados fora do país e apenas aplicados no Brasil (*outsourced bias*).

O PL 2338/2023 tenta mitigar esse problema ao exigir transparência e documentação de dados de treinamento, mas sem prever regras de rastreabilidade Inter jurisdicional, o que limita sua eficácia.

6.4. Erosão da confiança pública e o desafio da finalidade pública

A reutilização não controlada de dados e modelos mina a confiança pública em instituições e em pesquisas legítimas. Quando informações coletadas para fins de saúde pública ou ciência são posteriormente usadas para fins comerciais, há um desvirtuamento ético do propósito original. No Brasil, esse dilema é visível em casos como:

- o uso de dados do SUS por empresas privadas para desenvolver produtos de seguro ou marketing de saúde;
- a integração de bases governamentais sem consentimento explícito (como entre Receita Federal, Denatran e CadÚnico).

Isso compromete não só a confiança dos cidadãos, mas o próprio princípio constitucional da finalidade pública da administração (art. 37, caput, CF).

A solução exige uma nova regulação da finalidade dos modelos, alinhada ao interesse público e sujeita à fiscalização da ANPD e dos Tribunais de Contas.

6.5. Riscos Sistêmicos e a Centralidade da Supervisão Humana

18 BRASIL. PROJETO DE LEI Nº 2338/2023 – MARCO LEGAL DA

INTELIGÊNCIA ARTIFICIAL.

17 BRASIL. LEI Nº 13.709, DE 14 DE AGOSTO DE 2018. LEI GERAL DE PROTEÇÃO DE DADOS PESSOAIS (LGPD).

O principal perigo da inteligência artificial não está em uma possível "consciência das máquinas", mas na crescente complexificação dos sistemas sociais controlados por tecnologias sem compreensão da realidade. Erros algorítmicos em larga escala, em setores como energia, finanças, transporte ou segurança, podem gerar efeitos catastróficos.

No Brasil, o Conselho Nacional de Justiça (CNJ) já reconheceu o risco da chamada "automatização cega da Justiça" e, por meio das Resoluções nº 332/2020 e nº 455/2022, determinou que todo uso de IA no Poder Judiciário mantenha supervisão humana permanente, assegurando transparência e auditabilidade dos algoritmos. Esse mesmo princípio, conhecido como "human-in-the-loop", também está previsto no PL 2338/2023, que exige intervenção humana em sistemas de alto risco, em consonância com a proposta de Fowler, segundo a qual a técnica deve permanecer subordinada à inteligência sentiente, e não o contrário.

A discussão europeia sobre os limites do GDPR e as deficiências do AI Act evidencia desafios semelhantes aos enfrentados pelo Direito brasileiro. Embora a LGPD e o futuro Marco Legal da IA representem avanços importantes, ainda se concentram na proteção individual, deixando de abarcar de forma adequada os riscos sistêmicos e coletivos, especialmente os relacionados à reutilização de modelos e à assimetria informacional.

A proposta de "limitação da finalidade para modelos", presente na literatura internacional, pode inspirar um novo eixo regulatório nacional voltado à accountability retroativa, à supervisão pública e à proteção do interesse coletivo. Assim, o Direito brasileiro tem a oportunidade de avançar além do modelo europeu, adotando uma abordagem ética, preventiva e humanista, que reafirme a centralidade da pessoa humana e da verdade em meio à era da automação.

A diferença essencial entre o paradigma de conhecimento da inteligência artificial e o da inteligência humana senciente está em sua base epistemológica e em sua capacidade de perceber a realidade e conceber a verdade.

5. DA EPISTEMOLOGIA À REGULAÇÃO JURÍDICA DA IA

20 BRASIL. LEI Nº 13.709, DE 14 DE AGOSTO DE 2018. LEI GERAL DE PROTEÇÃO DE DADOS PESSOAIS (LGPD).

19 FOWLER, Thomas. Computers, Computation, and the Limits of Computability.

Washington, DC: XZF Technical Publications, 2021.

A distinção entre inteligência sensível e inteligência senciente ultrapassa o campo da filosofia e alcança implicações diretas para o Direito contemporâneo, especialmente no Brasil.

Enquanto a inteligência humana, por ser senciente, é capaz de compreender valores, contextos e a realidade moral das ações, a inteligência artificial, que se funda na sensibilidade empírica e na correlação estatística, ainda carece de intencionalidade, consciência e valor.

Essa diferença epistemológica justifica a necessidade de um regime jurídico próprio para o uso da IA, que reconheça sua utilidade instrumental, mas também imponha limites éticos e normativos. O Direito brasileiro, ancorado na dignidade da pessoa humana (art. 1º, III, CF), no devido processo legal (art. 5º, LIV e LV, CF) e na proteção de dados pessoais (Lei nº 13.709/2018 — LGPD), já fornece uma base sólida para essa limitação.

7.1. Explicabilidade e transparência

A limitação epistemológica da IA, ou seja, sua incapacidade de compreender a verdade se fundamenta juridicamente a exigência de explicabilidade dos algoritmos.

No Brasil, essa exigência é expressa nos arts. 6º, VI e X, da LGPD, que impõem os princípios da transparência e da responsabilização (*accountability*). O PL 2338/2023, em consonância com o *AI Act* europeu, prevê o dever de documentação e rastreabilidade das decisões automatizadas, garantindo ao cidadão o direito de compreender como e por que uma decisão foi produzida.

Essa transparência é indispensável para preservar o controle democrático e o contraditório, pilares do Estado de Direito.

7.2. Limitação da finalidade e proteção coletiva

As distinções filosóficas também iluminam o problema jurídico da reutilização secundária de modelos de IA. Como a IA carece de intencionalidade, é o legislador quem deve impor finalidade ética ao uso dos modelos.

22 BRASIL. PROJETO DE LEI Nº 2338/2023 – MARCO LEGAL DA INTELIGÊNCIA ARTIFICIAL.

No Brasil, o princípio da finalidade (art. 6º, I, LGPD) enfrenta o mesmo desafio que o GDPR europeu: a anonimização e o reaproveitamento de modelos tornam difícil controlar o propósito original dos dados. Por isso, parte da doutrina propõe que, ao adotar uma "limitação de finalidade para modelos", se vincule o uso da IA ao propósito legítimo e público para o qual foi desenvolvida.

Essa regulação seria coerente com o art. 37 da Constituição Federal, que consagra os princípios da finalidade, moralidade e publicidade na administração pública, e com a função social das tecnologias prevista no art. 170, III, CF.

7.3. Ética pública e accountability retroativa

Diante da assimetria informacional gerada por grandes modelos de IA, o Direito brasileiro deve fortalecer mecanismos de *accountability* retroativa, exigindo que desenvolvedores e operadores demonstrem a origem e o propósito dos dados utilizados no treinamento e a compatibilidade entre a finalidade original e a aplicação atual. Isso reforça a ideia de que o controle ético não é anterior apenas, é também permanente.

Em suma, as diferenças epistemológicas entre a inteligência sensível da IA e a inteligência senciente humana fornecem as bases filosóficas para as exigências jurídicas de supervisão humana, transparência e limitação de finalidade. Essas exigências não são meros instrumentos técnicos: são expressões jurídicas da consciência moral que distingue o humano do maquínico. No Direito brasileiro, esse princípio se traduz na máxima de que:

a automação deve servir à justiça, e não substituir o juízo.

7.4. Contratos, Acordos e Responsabilidade diferenciada

Diante desse cenário, torna-se essencial discutir a possibilidade de contratos com assistentes humanos, que reforçam a percepção de responsabilidade. Enquanto o assistente humano pode assumir obrigações contratuais — como cláusulas de confidencialidade, autoria compartilhada ou responsabilidade solidária —, a inteligência artificial não pode ser parte de contrato nem responder por inadimplemento.

Essa diferença cria uma assimetria institucional e perceptiva: o público tende a considerar o humano responsável porque o sistema jurídico o reconhece como sujeito de

direitos e deveres, enquanto a IA, sem personalidade jurídica nem vínculo contratual, é percebida como imune à punição, o que limita o reconhecimento de sua autoria moral ou criativa.

Dessa forma, a responsabilidade na era da IA não é apenas um tema tecnológico ou normativo, mas também cultural e cognitivo. O Direito deve reconhecer que a percepção social de autoria e culpa está em transformação e adaptar seus regimes de imputação. Compreender essa mudança é essencial para construir um modelo jurídico que, sem afastar a centralidade da pessoa humana, responda à crescente complexidade moral e operacional das inteligências assistidas.

6. REGIMES DE RESPONSABILIDADE E SOLUÇÕES REGULATÓRIAS

8.1. IA como Ferramenta e Responsabilidade Indireta

Na fase atual de desenvolvimento da IA, os robôs carecem de imputabilidade. O robô é ainda um produto da criação e manipulação humanas. A responsabilidade tende a recair sobre o ser humano que se apresenta como fabricante, operador, proprietário ou utilizador do robô.

A teoria da IA como ferramenta (*AI-as-tool*) vincula a responsabilidade objetiva pelo comportamento da máquina à pessoa em nome de quem ela age, com consequências similares à responsabilidade vicária ou pelo fato de terceiro.

Essa abordagem é consistente com o princípio de que a pessoa, natural ou jurídica, em cujo nome um computador foi programado, é responsável por qualquer mensagem gerada pela máquina (Pires & Silva, 2017). A responsabilidade do utilizador ou proprietário é ponderada em face do possível dever de vigilância ou guarda de coisa/animal. A Resolução do Parlamento Europeu (2015/2103(INL)) nota que a responsabilidade do agente humano deve ser proporcional ao nível de instruções dadas e ao grau de autonomia do robô (Pires & Silva, 2017).

8.2. IA como Produto e a Responsabilidade Objetiva do Fabricante

A imputação de danos ao fabricante é possível por meio do regime de responsabilidade civil do produtor, que se caracteriza como uma forma de

responsabilidade objetiva. Contudo, esse regime apresenta limitações, como a exigência de um defeito de fabrico do robô. A Diretiva Europeia 85/374/CEE, sobre responsabilidade por produtos defeituosos, pode ser aplicada, mas permanece a polêmica sobre a excludente de risco do desenvolvimento.

Nesse sentido, a Diretiva 85/374/CEE, em seu artigo 7º, alínea "e", dispõe que: *"O produtor não é responsável nos termos da presente directiva se provar: (...) e) Que o estado dos conhecimentos científicos e técnicos no momento da colocação em circulação do produto não lhe permitiu detectar a existência do defeito" (União Europeia, 1985).*

O problema reside em definir se os danos causados por atos independentes da IA, que age de forma não prevista por seus programadores devido à autoaprendizagem, estão abrangidos pelo risco do desenvolvimento ou se configuram, *res ipsa loquitor*, uma falha do produto. Reconhece-se que o atual quadro jurídico de responsabilidade pelo produto não é suficiente para abranger os danos provocados pela nova geração de robôs, devido às suas capacidades adaptativas e imprevisibilidade.

8.3. IA como Risco Criado: Responsabilidade Objetiva e Regulação Proposta

A expansão das novas tecnologias reforça a adoção de regimes de responsabilidade objetiva, estruturados em modelos econômicos de causalidade e esferas de risco (Antunes, 2024). Nesse regime, basta comprovar o dano e o nexo causal entre o funcionamento lesivo da IA e o prejuízo, dispensando-se a demonstração de culpa.

O Parlamento Europeu, em sua Resolução de 2017, reconheceu os desafios trazidos pela autonomia das máquinas e propôs soluções voltadas à tutela efetiva do lesado e à segurança jurídica dos produtores. Defendeu-se a responsabilidade objetiva fundada na gestão de riscos, atribuindo-a à pessoa mais apta a prevenir e mitigar danos, em lógica semelhante à teoria do bolso profundo (deep pocket) ‘’quem usufrui dos lucros da tecnologia deve também suportar seus riscos”. Entre as medidas propostas, destacam-se:

1. **Seguro Obrigatório de Responsabilidade Civil:** regime análogo ao da circulação de automóveis, impondo aos produtores ou proprietários de sistemas de IA a contratação de seguros para cobrir potenciais danos.
2. **Fundos de Compensação:** financiados por contribuições das empresas, atuariam como garantia para reparar danos não cobertos pelos seguros.
3. **Supervisão Humana Obrigatória (Human-in-the-loop):** a intervenção humana é exigida em sistemas de alto risco, assegurando discernimento ético e racional sobre decisões automatizadas. Esse princípio, que traduz a epistemologia da inteligência senciente em obrigação jurídica, encontra fundamento no art. 20 da LGPD (Lei nº 13.709/2018), que garante ao titular de dados o direito de revisão de decisões automatizadas, e é reforçado pelo PL 2338/2023 (Marco Legal da IA), que impõe supervisão humana contínua em áreas sensíveis como justiça, segurança pública e saúde.
4. **Accountability e Transparência:** necessidade de auditoria, documentação e rastreabilidade para garantir controle social e democrático sobre decisões automatizadas.

Apesar do debate sobre a atribuição de um estatuto de “pessoa eletrônica” aos robôs mais sofisticados, prevalece o entendimento de que, na ausência de personificação, a responsabilidade deve recair sobre um ser humano, seja com base no regime de produto defeituoso, de vigilância ou de risco. A supervisão humana, portanto, não é apenas requisito técnico, mas expressão da responsabilidade moral e jurídica indelegável do decisor humano frente à autonomia limitada das máquinas.

7. CONCLUSÃO

A reflexão sobre os desafios éticos e filosóficos da Inteligência Artificial no Direito contemporâneo revela que a questão central não está na tecnologia em si, mas na visão de conhecimento e de ser humano que a sustenta.

A IA, ao se basear em uma concepção empirista de inteligência sensível, é capaz de processar dados e reconhecer padrões, mas permanece incapaz de perceber a realidade, conceber a verdade ou exercer julgamento ético. O ser humano, ao contrário, é dotado de inteligência senciente, unidade viva entre razão, percepção e valor, que lhe permite compreender o mundo de modo criativo, moral e teleológico. No plano jurídico, essa distinção filosófica se traduz em obrigações normativas concretas.

A Constituição Federal brasileira, em seu art. 1º, III, coloca a dignidade da pessoa humana como fundamento da República. Esse princípio impõe que qualquer uso de

[27] BRASIL. Projeto de Lei nº 2338/2023 – Marco Legal da Inteligência Artificial.

tecnologias de decisão automatizada preserve o valor intrínseco da consciência e da autonomia humanas.

Em consonância, o art. 20 da Lei Geral de Proteção de Dados (Lei nº 13.709/2018) garante ao cidadão o direito à revisão humana de decisões automatizadas, reconhecendo que apenas a mente humana pode ponderar valores, interpretar contextos e agir com responsabilidade moral.

A futura aprovação do Marco Legal da Inteligência Artificial (PL 2338/2023) reforça esse horizonte ao prever supervisão humana obrigatória em sistemas de alto risco e ao consagrar princípios de transparência, explicabilidade e prevenção de danos coletivos.

Esses dispositivos não são meros requisitos técnicos: refletem, em linguagem jurídica, a intuição filosófica de que a máquina não conhece o real — apenas o simula. Por isso, o Direito deve garantir que o controle epistêmico e ético permaneça com o ser humano, sob pena de violar os fundamentos do Estado Democrático de Direito.

A IA, portanto, não ameaça o valor do humano; ela o reafirma. Seus limites ontológicos e epistemológicos revelam a irredutibilidade da inteligência humana, a única capaz de unir razão, emoção, valor e sentido. O Direito brasileiro, ao consolidar um modelo regulatório centrado na dignidade, na finalidade e na verdade, pode ocupar posição de vanguarda global ao traduzir em normas o que a filosofia já reconhece:

a justiça é um ato de inteligência senciente, e nenhuma máquina pode sentir a realidade que o Direito busca ordenar.

Assim, o desafio ético-jurídico contemporâneo não é conter o avanço da IA, mas reafirmar o primado do humano sobre a técnica. A tecnologia deve ser instrumento de

libertação, não de alienação; meio de aperfeiçoar o juízo, não de substituí-lo. O futuro do Direito, e da própria democracia, dependerá da nossa capacidade de manter o ser humano no centro das decisões, não como peça da engrenagem algorítmica, mas como sujeito moral e criador de sentido.

REFERÊNCIAS

ANTUNES, Henrique Sousa. **Inteligência Artificial e responsabilidade civil: enquadramento.** *In*: ESTUDOS DE HOMENAGEM AO SENHOR DR. RUI PENA. [S.l.: s.n.], 2024.

BRASIL. **Constituição Federal de 1988.**

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. **Lei Geral de Proteção de Dados Pessoais (LGPD).**

BRASIL. Projeto de Lei nº 2338/2023 – **Marco Legal da Inteligência Artificial.**
ČERKA, Paulius; GRIGIENĖ, Jurgita; SIRBIKYTĖ, Gintarė. **Liability for damages caused by Artificial Intelligence.** *Computer Law & Security Review*, Elsevier,

v. 31, n. 3, p. 376-389, jun. 2015.

FOWLER, Thomas. **Computers, Computation, and the Limits of Computability.** Washington, DC: XZF Technical Publications, 2021.

HUME, David. **An Enquiry Concerning Human Understanding.** 1748. HUME, David. **Treatise of Human Understanding.** Oxford: Oxford University

Press, 1951.

PAGALLO, Ugo. **The laws of robots: crimes, contracts, and torts.** Heidelberg: Springer, 2013.

PIRES, Thatiane Cristina Fontão; SILVA, Rafael Peteffi da. **A responsabilidade civil pelos atos autônomos da inteligência artificial: notas iniciais sobre a resolução do Parlamento Europeu.** *Revista Brasileira de Políticas Públicas*, Brasília, v. 7, n. 3, p. 238-254, 2017.

RUSSEL, Stuart; NORVING, Peter. **Artificial Intelligence: a modern approach.** 3. ed. Harlow (UK): Pearson Education Limited, 2014.

UNIÃO EUROPEIA. Diretiva 85/374/CEE do Conselho, de 25 de julho de 1985. UNIÃO EUROPEIA. **Resolução do Parlamento Europeu**, de 16 de fevereiro de

2017, com recomendações à Comissão de Direito Civil sobre Robótica (2015/2103(INL)). 2017.

ZUBIRI, Xavier. **God and Deification in Pauline Theology.** Lanham, MD: University Press of America, 1981.