

CAROLINE DA ROSA CAVALHEIRO

PREFÁCIO DE Cristiano Becker Isaia

SISTEMAS DE COMPORTAMENTO INTELIGENTE E NEOLIBERALISMO PROCESSUAL

A ARQUITETURA DA MATRIX E O DIREITO À
EXPLICABILIDADE “POST-HOC”




EDITORA
ILUSTRAÇÃO

CAROLINE DA ROSA CAVALHEIRO

**SISTEMAS DE COMPORTAMENTO
INTELIGENTE E NEOLIBERALISMO PROCESSUAL**

A ARQUITETURA DA MATRIX E O DIREITO À
EXPLICABILIDADE “Post-Hoc”

Editora Ilustração
Santo Ângelo – Brasil
2026



Esta obra está licenciada com uma Licença Creative Commons

<https://creativecommons.org/licenses/by-nc-sa/4.0>

Capa: Freepik e IA

Revisão: A autora

CATALOGAÇÃO NA FONTE

C376s Cavalheiro, Caroline da Rosa

Sistemas de comportamento inteligente e neoliberalismo processual [recurso eletrônico] : a arquitetura da Matrix e o direito à explicabilidade “post-hoc” / Caroline Cavalheiro. - Santo Ângelo : Ilustração, 2026.

183 p.

ISBN 978-65-6135-214-7

DOI 10.46550/978-65-6135-214-7

1. Direito. 2. Inteligência artificial. 3. Inovação digital I.
Título

CDU: 34:004

Responsável pela catalogação: Fernanda Ribeiro Paz - CRB 10/ 1720



Crossref



E-mail: eilustracao@gmail.com

www.editorailustracao.com.br

Conselho Editorial



Dra. Adriana Maria Andreis	UFFS, Chapecó, SC, Brasil
Dra. Adriana Mattar Maamari	UFSCAR, São Carlos, SP, Brasil
Dra. Berenice Beatriz Rossner Whatuba	URI, Santo Ângelo, RS, Brasil
Dr. Clemente Herrero Fabregat	UAM, Madri, Espanha
Dr. Daniel Vindas Sánchez	UNA, San Jose, Costa Rica
Dra. Denise Tatiane Girardon dos Santos	UNICRUZ, Cruz Alta, RS, Brasil
Dr. Domingos Benedetti Rodrigues	UNICRUZ, Cruz Alta, RS, Brasil
Dr. Edeimar Rotta	UFFS, Cerro Largo, RS, Brasil
Dr. Edivaldo José Bortoleto	UNOCHAPECÓ, Chapecó, SC, Brasil
Dra. Elizabeth Fontoura Dorneles	UNICRUZ, Cruz Alta, RS, Brasil
Dr. Evaldo Becker	UFS, São Cristóvão, SE, Brasil
Dr. Glaucio Bezerra Brandão	UFRN, Natal, RN, Brasil
Dr. Gonzalo Salerno	UNCA, Catamarca, Argentina
Dr. Héctor V. Castanheda Midence	USAC, Guatemala
Dr. José Pedro Boufleuer	UNIJUÍ, Ijuí, RS, Brasil
Dra. Keiciane C. Drehmer-Marques	UFSC, Florianópolis, RS, Brasil
Dr. Luiz Augusto Passos	UFMT, Cuiabá, MT, Brasil
Dra. Maria Cristina Leandro Ferreira	UFRGS, Porto Alegre, RS, Brasil
Dra. Neusa Maria John Scheid	URI, Santo Ângelo, RS, Brasil
Dra. Odete Maria de Oliveira	UNOCHAPECÓ, Chapecó, SC, Brasil
Dra. Rosângela Angelin	URI, Santo Ângelo, RS, Brasil
Dr. Roque Ismael da Costa Güllich	UFFS, Cerro Largo, RS, Brasil
Dra. Salete Oro Boff	ATITUS, Passo Fundo, RS, Brasil
Dr. Tiago Anderson Brutti	UNICRUZ, Cruz Alta, RS, Brasil
Dr. Vantoir Roberto Brancher	IFFAR, Santa Maria, RS, Brasil

Este livro foi avaliado e aprovado por pareceristas *ad hoc*.

O presente livro é fruto da Dissertação de Mestrado realizada por Caroline da Rosa Cavalheiro, apresentada ao Programa de Pós Graduação em Direito da Universidade Federal de Santa Maria (UFSM/RS), sob a área de concentração “Direitos Emergentes na Sociedade Global”, com ênfase na Linha de Pesquisa “Direitos na sociedade em rede: atores, fatores e processos na mundialização”, como requisito parcial para obtenção do título de Mestre em Direito.

A pesquisa foi elaborada sob a orientação do Professor pós-Doutor em Direito Público Cristiano Becker Isaia, o qual com maestria “amadrinhou” a discente com generosidade e firmeza. Existem reflexões que permanecerão além destas páginas e do período de convívio. Ademais, contou-se com a coorientação do Professor Doutor Alexandre de Oliveira Zamberlan, cuja presença foi fundamental para que fosse aberta a cancela que conduz ao caminho da ciência da computação. Com suas orientações muitos horizontes se descortinaram.

A autora dedica esta obra aos seus pais, Sandra da Rosa Cavalheiro e José Machado Cavalheiro (*in memoriam*), pelo amor incondicional e pelos ensinamentos de como ser e viver. Agradece por serem seu esteio e por mostrarem que, com dedicação e humildade, não existe impossível.

Cada um sabe seu dom na jornada que empreende e pelo
tempo entende o quanto o amor nos ampara. Somente o
tempo não para, recria-se e permanece. Na intensidade da
prece que de esperança gravita, pois quando a fé nos habita é
que o milagre acontece.

(Gujo Teixeira, 2025)

SUMÁRIO

PREFÁCIO.....	15
1 INTRODUÇÃO	19
2 A INTELIGÊNCIA ARTIFICIAL DIANTE DA EXPLORAÇÃO DO SUBSTRATO DOS SISTEMAS DE COMPORTAMENTO INTELIGENTE	25
2.1 Navegando com a Nabucodonosor pelos algoritmos: investigando as dimensões do sistema de comportamento inteligente.....	31
2.2 Reestruturando a matrix: principais regulamentações da IA como recarga do sistema	56
3 O CENÁRIO ESTABELECIDO PELO PROCESSUALISMO (TECNO)NEOLIBERAL E OS DESAFIOS DA INTELIGÊNCIA ARTIFICIAL	75
3.1 Bullet Time e a virada tecnológica no poder judiciário e(m) crise ...	83
3.2 Metacortex, automação processual e estratégias de modernização e eficiência operacional: o exemplo privilegiado do IRDR, sistema de precedentes	95
4 INTELIGÊNCIA ARTIFICIAL E DECISÃO JUDICIAL: RISCOS E BENEFÍCIOS NO USO DA IA	115
4.1 O déjà vu a (des)serviço do direito: a implementação da IA no poder judiciário	124
4.2 Re(arquitetando) o direito à explicabilidade “post-hoc” como forma de decodificar os limites do código em busca da resposta correta: heurísticas, vieses e opacidades dos sistemas de comportamento inteligente na tomada de decisão jurisdicional	137
5 CONCLUSÃO	161
REFERÊNCIAS.....	169
SOBRE A AUTORA	183

PREFÁCIO

A obra que se está a prefaciар é resultado de uma pesquisa marcada pela profundidade teórica, metodológica, e pelo rigor intelectual. Fruto da Dissertação de Mestrado desenvolvida por Caroline da Rosa Cavalheiro no Programa de Pós-Graduação em Direito da Universidade Federal de Santa Maria, este livro revela uma jornada investigativa invulgar, curiosidade científica e dedicação incansável ao estudo crítico da relação entre inteligência artificial, sistemas de comportamento inteligente e o chamado neoliberalismo processual.

A autora percorreu um caminho de pesquisa que não se limitou à tradicional revisão bibliográfica. Sua investigação exigiu atravessar fronteiras epistemológicas, dialogar com campos diversos — como filosofia, ciência da computação, neurociência, história da tecnologia, hermenêutica jurídica e teoria do processo — e articular perspectivas que, à primeira vista, poderiam parecer díspares. Esse gesto intelectual revela não apenas esforço, mas também coragem acadêmica.

Neste trabalho, Caroline Cavalheiro busca compreender, a partir do fenômeno contemporâneo da inteligência artificial, as transformações profundas que atravessam o Direito, em especial o campo processual. A pesquisa parte de uma constatação que se impõe: a crescente integração de sistemas automatizados, algoritmos decisórios e plataformas digitais vêm modificando de modo radical o funcionamento do Poder Judiciário. Longe de adotar uma perspectiva simplesmente entusiasmada ou ingênua, a autora investiga esse cenário com prudência, oferecendo uma reflexão sólida sobre limites, riscos, benefícios e tensões decorrentes dessa virada tecnológica.

A atualidade da temática é inegável. Nunca se falou tanto sobre inteligência artificial, automação e algoritmização. Entretanto, são poucos os estudos que ultrapassam a superfície do debate e se dedicam a investigar, de forma rigorosa, o substrato técnico dos sistemas de comportamento inteligente e suas implicações normativas. Este livro se distingue justamente por enfrentar essa lacuna, traduzindo conceitos computacionais complexos em linguagem acessível, sem jamais perder densidade teórica ou precisão conceitual.

Há, ao longo da obra, um esforço visível de desconstruir simplificações e iluminar zonas de sombra deixadas por discursos tecnocráticos. A autora examina a estrutura técnica dos algoritmos, discute a lógica dos sistemas de classificação, aborda fenômenos como aprendizado de máquina e heurísticas automatizadas, e revela como essas dinâmicas podem influenciar (e por vezes enviesar) a tomada de decisão judicial.

Seu estudo vai além dos efeitos imediatos e olha para o fenômeno de modo estrutural, investigando como a lógica gerencialista, característica do processualismo tecnoneoliberal, tende a moldar, padronizar e acelerar decisões que deveriam permanecer ancoradas em princípios constitucionais, na racionalidade jurídica e na integridade interpretativa.

O mérito da obra também está na forma como integra referências culturais e filosóficas, utilizando metáforas como Matrix e a alegoria da caverna de Platão não como artifícios retóricos, mas como instrumentos pedagógicos que ajudam a compreender a complexidade das relações entre informação, controle, realidade construída e ilusão tecnológica. Essas analogias, longe de simplificarem o conteúdo, aprofundam a reflexão e tornam o texto instigante, sem perder rigor acadêmico.

Não menos importante é reconhecer o empenho pessoal que permeou toda a trajetória de pesquisa. Caroline Cavalheiro demonstra, em cada capítulo, que não se trata apenas de um estudo técnico, mas de um exercício de responsabilidade científica e social. Sua investigação nasce do desejo de compreender para transformar; de diagnosticar para proteger; de aprender para, quem sabe, desvelar caminhos futuros. A autora estudou longamente, revisou incessantemente, ajustou conceitos, reescreveu argumentos e retornou às fontes sempre que necessário. Essa persistência reverbera no texto e confere à obra a solidez que o leitor encontrará nas próximas páginas.

Assim, este livro oferece ao público especializado uma reflexão indispensável para compreender os desafios éticos, epistemológicos e institucionais da era da inteligência artificial. Seu valor reside tanto na atualidade do tema quanto na profundidade com que é tratado.

Em suma, esta obra reflete o comprometimento da autora com a pesquisa séria, honesta e intelectualmente generosa. É resultado de um trabalho intenso, perseverante e guiado por um profundo senso de responsabilidade acadêmica e social. Que sua leitura inspire novos debates, provoque novas investigações e contribua para o fortalecimento de um

Direito capaz de enfrentar, com integridade, os desafios trazidos pelos sistemas de comportamento inteligente e pela tecnicização crescente de nossas instituições.

Prof. Dr. Cristiano Becker Isaia

Universidade Federal de Santa Maria Programa de Pós-Graduação
em Direito (Mestrado e Doutorado)

1 INTRODUÇÃO

“What do I got to do to wake you up? To shake ya up, to break the structure up?” (Rage Against the Machine, 1992)

A provocação lançada pela banda *Rage Against the Machine* em *Wake Up* funciona como um grito contra sistemas de dominação que se perpetuam sob o manto da normalidade. Ao questionar estruturas estabelecidas, a letra convida ao despertar crítico, ecoando a alegoria de *Matrix* (Wachowski; Wachowski, 1999), onde a realidade é uma construção artificial que oculta os verdadeiros mecanismos de controle. Assim como no filme, a música incita o rompimento com padrões impostos e conduz à busca por autonomia diante de forças invisíveis, tema que se torna especialmente relevante diante da ascensão das tecnologias inteligentes.

A inteligência artificial (IA) alcançou notável protagonismo no século XXI, notadamente a partir do ano de 2022, com o advento de sistemas dotados de modelos de linguagem avançados, aptos a produzir textos, oferecer respostas com elevada precisão e auxiliar em múltiplas tarefas cotidianas. A inovação transcendeu o interesse dos entusiastas da tecnologia, consolidando-se como ferramenta indispensável em diversos contextos profissionais e ambientes laborais.

Todavia, a IA não surgiu de forma abrupta, visto que é fruto de um longo processo de abstração, em que se fragmentam dimensões essenciais da experiência humana: emoção e razão, símbolo e número, memória e processamento. Seu quadro histórico se arremanga em 1640 e abre passagem para divisões que atravessam as ciências exatas, naturais e humanas. Há um rompimento entre duas formas clássicas de compreender o mundo: o mecanismo que vê a realidade como um conjunto de engrenagens regidas por leis físicas e a teleologia que a entende como orientada a um fim, como Aristóteles acreditava.

1 Em tradução livre pela autora: “O que eu tenho que fazer, o que eu tenho que fazer para te acordar? Para te chacoalhar, para quebrar a estrutura?”. Citação referente à parte da canção *Wake up*, da banda de Zack de la Rocha, que toca nos créditos de encerramento do primeiro filme da trilogia de *Matrix*, de 1999. O trecho da música usa o verbo “acordar” não apenas como um despertar literal, mas como um despertar para a realidade social e política, incitando o ouvinte a questionar a autoridade, a despertar sua consciência e a se juntar à luta por mudança.

Aquilo que em Aristóteles era potência e ato, que em Kant era razão pura e razão prática, torna-se, no século XX, memória e processamento, heurística e algoritmo. A inteligência artificial é um processo civilizatório que busca transformar o pensar profundo em raciocinar de forma matemática, rápida e eficiente. Afasta-se cada vez mais da ideia de propósito (sentido) e se aproxima da ideia de funcionalidade (eficiência). Desse modo, Manuel Castells contribui com o conceito de sociedade em rede, expressão que designa uma nova morfologia social constituída por fluxos de informação instantâneos e descentralizados. E Lênio Streck complementa ao falar com sapiência que a era da técnica (*Ges-tell*) e do virtual são duas dimensões de um mesmo tempo.

O “*homo technologicus*” inserido em uma lógica de rede impõe uma reorganização profunda na comunidade jurídica. Na medida em que a historicidade do processo judicial remonta ao período clássico romano, com destaque para o período justiniano, quando emergiu o rito ordinário e se delineou a distinção entre julgamento e decisão. Posteriormente, as correntes filosóficas racionais e iluministas sedimentaram essa visão ao sustentar que a sentença, que encerra o processo, é proferida após uma exaustiva atividade cognitiva, o que a capacita, por essa razão, a desvendar o verdadeiro “sentido da lei”. Nesse sentido, esse impasse estrutural é a causa da morosidade do sistema jurídico em salvaguardar a efetivação de direitos e a IA é vista como possibilidade de cura.

Com isso, o Poder Judiciário brasileiro vive uma verdadeira “virada tecnológica”, que se materializa em uma de suas fases por meio da integração de ferramentas de IA na rotina decisória e administrativa. Atualmente, o judiciário conta com 147 sistemas de inteligência artificial integrados à plataforma Sinapses. No Supremo Tribunal Federal, destacam-se os robôs Victor, Rafa, VitórIA e MarIA, voltados à triagem e organização processual, bem como à elaboração de documentos. O Superior Tribunal de Justiça conta com o Athos, que detecta documentos correlatos, enquanto o Supremo Tribunal Militar emprega o Lara para simplificação textual.

Diversos Tribunais Regionais e Estaduais implementaram soluções que automatizam etapas como agrupamento de processos, confecção de minutas, transcrição de audiências. Em algumas regiões, a escolha das nomenclaturas incorporam referências culturais significativas como é o caso do robô Oxóssi, que remete à espiritualidade afro-brasileira; do Jerimum, evocando a identidade nordestina; e do Maria Firmina, que homenageia a primeira romancista negra do Brasil. Outras denominações, como Arandu

que advém do termo tupi que significa sabedoria e Toth inspirado no deus egípcio do conhecimento, conferindo a esses sistemas um caráter simbólico da riqueza cultural brasileira.

Ainda que as facilidades oferecidas pela inteligência artificial sejam sedutoras nesse mundo líquido, como bem observa Bauman, surgem anseios legítimos por segurança normativa, transparência e resguardo ético em meio à vertigem das inovações do campo do jurídico. Ao passo que, também, a interpretação do direito e dos casos exige um profundo conhecimento do contexto social, histórico e cultural. Dessa forma, a presente pesquisa será focada no estudo da tomada de decisão estabelecida pelos sistemas de comportamento inteligente, especialmente no ordenamento jurídico brasileiro na atualidade. Assim, diante da relação entre processo judicial, decisão e tecnologia, quais os limites e possibilidades para se equilibrar os riscos e benefícios da inteligência artificial no campo do direito?

Como luzeiro em uma noite cerrada, o objetivo geral consiste em compreender de forma crítico-reflexiva as nuances intrínsecas do uso da inteligência artificial no Judiciário brasileiro, particularmente no que se refere aos riscos e benefícios associados à tomada de decisão alicerçada em sistemas de comportamento inteligente. Já, os escopos específicos são: apresentar aportes teóricos iniciais sobre a inteligência artificial; contextualizar a sociedade em rede e a virada tecnológica no âmbito jurídico do Brasil; e abordar, sob a ótica hermenêutica, a respeito dos riscos e benefícios na tomada de decisão brasileira albergada pelos sistemas de comportamento inteligente.

A relevância desta pesquisa reside, socialmente, na imperiosa necessidade de desvendar o funcionamento da compreensão, interpretação e aplicação da lei em um contexto onde sistemas de comportamento inteligente começam a influenciar decisivamente a tomada de decisão judicial. Levando em conta que a virada tecnológica, catalisada pela pandemia, acelerou a digitalização e automação dos processos judiciais, refletindo uma mudança paradigmática no campo jurídico.

Seja dito de passagem que o direito, por sua própria natureza, já possui uma linguagem de difícil compreensão e somado à utilização de robôs e demais sistemas automatizados, torna-se ainda mais angustiante para os integrantes da sociedade entenderem como seus direitos são assegurados. Soma-se a este contexto, a incessante busca por eficiência e celeridade, que muitas vezes prioriza a quantidade em detrimento da

qualidade das decisões, criando um processo tecnoliberal que transforma o direito em um mecanismo de produção em massa de respostas processuais.

Revela-se como um terreno pedregoso e desafiador para os operadores do direito cujos conhecimentos jurídicos, por mais profundos que sejam, não bastam para enfrentar as complexidades intrínsecas à linguagem e ao funcionamento dessas novas tecnologias. Ou seja, a situação é ainda mais complexa devido à necessidade de compreender o funcionamento das máquinas, os algoritmos subjacentes e os comandos que orientam suas operações para que não se caia no limbo de ilusão de escolha (decisão manipulada pela máquina).

Por isso, é elementar que os profissionais do direito se familiarizem, por exemplo, com conceitos de termos técnicos como “*input*” e “*output*”, aprendizado de máquina, aprendizado profundo e redes neurais. Adicionalmente, há uma necessidade de ler textos que frequentemente provêm do inglês, o que requer não apenas a assimilação de novos conhecimentos, mas também a integração de ensinamentos estrangeiros. Dado que esse tema evoca o legado do britânico Alan Turing, o qual é reconhecido mundialmente como o pai da ciência da computação moderna e da inteligência artificial.

O panorama supracitado impõe um desafio significativo, que transcende os limites da expertise jurídica convencional e demanda uma interseção de conhecimentos multidisciplinares, onde o direito e a inovação digital se entrelaçam formando uma complexa sociedade em rede (de direitos e tecnologias). É nessa união de saberes que se encontra o potencial para desbravar esse novo território, assegurando que a justiça, mesmo diante de inovações tão profundas, continue a ser um pilar de equidade e transparência para todos. Desse modo, o presente estudo, será de grande valia para todos cidadãos que aspiram por um sistema judicial justo, eficiente e coerente.

Por seu turno, cientificamente, à luz da hermenêutica, será possível um entendimento mais aprofundado sobre como garantir a obtenção de resposta constitucionalmente adequada e sem arbitrariedades na atuação judicial automatizada. Considera-se que os vieses cognitivos, dos programadores, podem se refletir nos sistemas de comportamento inteligente e, conseqüentemente, nas decisões automatizadas, perpetuando desigualdades e injustiças.

Dessa forma, surge a curial importância do direito à explicabilidade *post-hoc*, que consiste na capacidade de explicar, após o fato, como uma

decisão foi tomada pelo sistema (seja em casos fáceis ou difíceis). Isso pode ser feito através de elucidações por textos, imagens, pesos de variáveis ou resumos de processos, descomplicando e clarificando o funcionamento dos códigos. Permitindo não apenas a verificação da legalidade das decisões automatizadas, mas também a correção de possíveis erros ou distorções, assegurando a proteção dos direitos fundamentais e um julgamento íntegro e confiável em um ambiente cada vez mais tecnológico.

Para além dessas considerações, perquire-se se a utilização de “métodos” automatizados no campo jurídico não apenas sugere uma regressão, mas também uma ilusão de progresso, uma falsa transição de meios que repete ciclos históricos ao tentar converter o direito, uma ciência do espírito, em uma ciência exata. Ao submeter a aplicação da lei a algoritmos, como se fossem meros postulados matemáticos, correndo-se o risco de despojar o direito de sua essência humanística, a uma redução de questões jurídicas a equações frias. Essa mecanização pode desconstruir a riqueza interpretativa e ética que é inerente ao direito, desviando-o de sua verdadeira natureza e função social.

Para mais, não se pode ignorar a possibilidade do esquecimento ou desmerecimento da linguagem como veículo primordial de expressão do pensamento jurídico. A substituição do discurso humano por respostas automatizadas não apenas empobrece o debate, mas também desvirtua a profundidade e a nuance que só a linguagem pode captar, ameaçando reduzir o direito a um conjunto de fórmulas desprovidas de alma.

Ao fornecer *insights* embasados nessas abordagens, este estudo se conecta diretamente às investigações realizadas pela pesquisadora tanto na graduação quanto no decorrer do mestrado, ambas com foco nas alterações processuais advindas da tecnologia, alinhadas à linha de pesquisa de direitos na sociedade em rede. Isso se evidencia pelo fato de que durante a graduação, em 2022, o tema de seu trabalho de conclusão de curso consistiu nas implicações da Lei Geral de Proteção de Dados (LGPD) no Processo Judicial Eletrônico (PJE), e, após dois anos, as investigações prosseguem na mesma linhagem de processo e tecnologia.

No entanto, em tão pouco tempo, o processo foi progressivamente dominado pela inteligência artificial, por novas legislações e por novas preocupações emergentes. Esse cenário levanta indagações sobre a natureza da era em que estamos inseridos, evocando a metáfora do filme *Matrix*, onde a arquitetura de nossa realidade social está cada vez mais sendo moldada pelas redes digitais. Assim, como na metáfora da caverna de

Platão, os juristas são confrontados com escolhas que, simbolicamente, se assemelham à decisão entre a pílula vermelha e a pílula azul.

Essas escolhas refletem não apenas a dualidade entre a aceitação de uma realidade ilusória ou o enfrentamento de uma verdade desconfortável. Isto é, na narrativa platônica, a caverna simboliza o mundo sensível, onde as sombras projetadas nas paredes representam as percepções enganosas baseadas nos sentidos limitados dos prisioneiros. Em contraste, a luz do sol, que está no exterior, simboliza o mundo das ideias ou o conhecimento verdadeiro, que só pode ser alcançado através da filosofia.

No filme supramencionado, o personagem Oráculo revela que a essência do comportamento humano repousa nas escolhas e não na mera binariedade de causa e efeito. A escolha de Neo em engolir a pílula vermelha e posteriormente rejeitar a reinserção do código na fonte, que permitiria a (re)inicialização da Matrix, torna-se um símbolo da busca pela verdade além das ilusões, representando a aspiração por um entendimento mais profundo e genuíno da realidade. Portanto, a busca pela “luz” através da presente investigação, sobre como funciona a arquitetura tecnológica que está moldando a sociedade, poderá evitar que as sombras projetadas por sistemas automatizados sejam aceitas como a única realidade.

Para tanto, utilizar-se-á como teoria de base a fenomenologia-hermenêutica, promovendo uma compreensão profunda a respeito desse novo mundo configurado pela tecnicização. Em termos de procedimento, empregar-se-á o bibliográfico e documental, com técnicas de fichamentos e resumos, com vistas a explorar normativas e reflexões teóricas produzidas sobre a matéria.

2 A INTELIGÊNCIA ARTIFICIAL DIANTE DA EXPLORAÇÃO DO SUBSTRATO DOS SISTEMAS DE COMPORTAMENTO INTELIGENTE

“Bem-vindo ao deserto do real.”

– Morpheus, em *Matrix* (Wachowski; Wachowski, 1999)

Foi em uma manhã de reflexões e ajustes no projeto, no campus da universidade, durante uma conversa descontraída com o orientador que emergiu a ideia de estabelecer um elo entre a temática da pesquisa e o universo cinematográfico de *Matrix*. Essa conexão encontrou solo fértil ao confrontar os dilemas tecnológicos e filosóficos presentes no filme com os desafios jurídicos e contemporâneos que permeiam a presente pesquisa.

A narrativa ao explorar a tensão entre realidade e ilusão, revela-se uma metáfora provocativa, convidando a desvendar as estruturas invisíveis de controle que moldam não apenas o sistema, mas também a experiência daqueles que estão imersos nele. Dessa forma, encontra-se ressonância no campo jurídico onde a inteligência artificial coloca em xeque os alicerces éticos, sociais e institucionais da justiça. Propõe-se ao leitor uma jornada reflexiva sobre a possibilidade de vislumbrar um futuro jurídico que conjugue inovação com equidade.

Acredita-se que para compreender as circunstâncias do presente é de curial importância realizar o acesso ao disco rígido (HD) e ao hipocampo para que as informações da história sejam resgatadas. Ambos são fundamentais para a preservação de experiências e a recuperação de informações quando necessário, permitindo que o indivíduo ou o sistema acesse essas memórias em diferentes contextos.

O acesso a essas “bibliotecas” é essencial para a construção de uma identidade coesa e para o entendimento das dinâmicas sociais que permeiam a contemporaneidade. Compreender o presente exige escavar o passado. Afinal, a ciência é um fenômeno histórico por excelência, embora busque produzir a cada instante uma ciência atemporal, como leciona Allen Newell (Newell, 1983). A ciência edificou seu principal sistema de

recompensa em torno da descoberta e da invenção de noções que são, em sua essência, históricas (Newell, 1983). Isto é, as grandes invenções não surgem do nada, pois estão ancoradas em contextos já existentes.

Desse modo, como uma paixão que beira a compulsão, acolhe a advertência de que ignorar o passado é estar destinado a repeti-lo (Newell, 1983). Independente do tempo, os equívocos e descobertas pretéritas são elementos essenciais para a evolução da sapiência (*phronesis*) e ao (re)visitá-los há a possibilidade de esculpir “melhor” o futuro do conhecimento científico. Deveras, a sabedoria prática deve, portanto, ser uma verdadeira e ponderada capacidade de agir com respeito aos bens humanos (Aristóteles, 1991).

O objeto do conhecimento científico existe de forma necessária, o que implica em sua eternidade, já que tudo aquilo que existe por necessidade em sentido absoluto é eterno, sendo as coisas eternas não geradas e imperecíveis (Aristóteles, 1991). Enquanto a primeira está voltada para o agir no mundo humano e mutável, o segundo lida com verdades que são necessárias, imutáveis e por isso consideradas eternas. Entra-se no túnel do tempo levando em conta que nosso mundo se encontra imerso em um processo de transformação estrutural que se estende por duas décadas (Castells; Cardoso 2005). A partir de 1990 a revolução da informação e a disseminação das tecnologias digitais transformaram a forma como as pessoas se comunicam, interagem e organizam socialmente.

Embora o surgimento da Internet remonte à década de 1960, seu uso se consolidou como uma tendência predominante nesse período. Dado que a sociedade molda a tecnologia conforme as necessidades, valores e interesses daqueles que a utilizam (Castells; Cardoso 2005). Diga-se de passagem que em 1954, Allen Newell e Herbert Simon foram os pioneiros a iniciar a programação de computadores voltados para a inteligência geral (McCarthy, 2007). Iniciou-se a ousada empreitada de insuflar pensamento às máquinas, na esperança de que elas pudessem aprender e resolver problemas como verdadeiros espelhos da mente humana.

Aliás, menciona-se que na cibernética, além da teoria da transmissão de mensagens na engenharia elétrica, abrange-se um campo mais amplo que incorpora não só o estudo da linguagem, mas também o estudo das mensagens como ferramentas para orientar a maquinaria e a sociedade, o desenvolvimento de computadores e outros autômatos, reflexões específicas sobre psicologia e sistema nervoso, além de uma nova teoria conjectural sobre o método científico (Wiener, 1954).

Trata-se de um estudo transdisciplinar, que rompe com o modelo linear tradicional de produção do conhecimento, integrando diversas áreas do saber que compõem a complexidade da sociedade humana. Aquilo que hoje denominamos neurociência já vinha sendo investigado sob outras abordagens há séculos.

Pode-se argumentar que, atualmente, a saúde, o poder e a geração de conhecimento dependem em grande medida da capacidade de organizar a sociedade para aproveitar os benefícios do novo sistema tecnológico, fundamentado na microeletrônica, nos computadores e na comunicação digital, com uma conexão cada vez mais intensa com a revolução biológica e suas ramificações, como a engenharia genética (Castells; Cardoso, 2005). A contemporaneidade presencia o surgimento de inquietações jurídicas inéditas dentre as quais se destaca o resguardo jurídico dos denominados neurodireitos no intuito de salvaguardar inviolabilidade da mente humana em virtude do avanço da neurotecnologia.

É oportuno assinalar que, como sublinha Castells, os vocábulos “informação” e “comunicação” não são criações da modernidade, pois sempre acompanharam a trajetória histórica da humanidade. O que se apresenta como novo é o aparato tecnológico que dá suporte à vida social. A verdadeira inovação reside no emprego da microeletrônica como base estruturante, a partir de redes tecnológicas que conferem novas potencialidades a uma forma ancestral de organização social: as redes (Castells; Cardoso, 2005).

As redes de comunicação digital constituem a coluna vertebral da sociedade em rede, assim como as redes de potência (ou redes energéticas) que formaram as infraestruturas que sustentaram a sociedade industrial, conforme demonstrado pelo historiador Thomas Hughes (Castells; Cardoso, 2005). Enquanto a era industrial dependia de uma infraestrutura física centrada em energia e produção, a da sociedade em rede repousa sobre uma infraestrutura imaterial. De fato, a sociedade em rede se manifesta de diversas maneiras, influenciada pela cultura, pelas instituições e pela trajetória histórica de cada sociedade, tal como como a sociedade industrial abrangeu realidades tão distintas quanto os Estados Unidos e a União Soviética, ou a Inglaterra e o Japão, essas sociedades compartilhavam características fundamentais que permitiam sua definição dentro do industrialismo, como forma de organização humana não era determinada pelas tecnologias industriais, mas se tornava impensável sem a presença delas (Castells; Cardoso, 2005).

Então, observa-se que a placa mãe que integra e conecta todos os componentes essenciais do sistema, como o processador, a memória e os dispositivos de armazenamento, serve como estrutura central do corpo social sustentado pelas conexões comunicacionais digitais. Ao passo que as redes interativas computacionais proliferam de forma exponencial, engendrando novas modalidades e canais de comunicação que não apenas exercem uma influência marcante sobre o curso da vida, mas também se deixam moldar por ela, em um processo simbiótico de transformação recíproca e contínua (Castells, 1999).

A representação da cidade contemporânea não se define mais pelo cerimonial da abertura dos portões, pelo ritual das procissões ou pela sequência ordenada de ruas e avenidas; a arquitetura urbana, agora, deve se vincular à criação de um “espaço-tempo tecnológico” (Virilio, 1993). Isso significa que a cidade da modernidade é moldada pela velocidade e pela tecnologia, substituindo a organização física e tradicional por uma fluidez que conecta o espaço urbano ao tempo digital.

Com isso, o protocolo de acesso da telemática substitui o do portão e aos tambores das portas se sucedem os do bancos de dados (Virilio, 1993). O que dá origem a perquirição de “Arquitetura ou pós-arquitetura? Definitivamente o debate em torno da modernidade parece participar de um fenômeno de desrealização que atinge, de uma só vez, as disciplinas de expressão, as formas de representação e informação (Virilio, 1993). O “portão” representa uma antiga limitação física, um ponto de passagem ou defesa visível. Consequentemente, com o avanço da integração entre telecomunicações e informática (telemática), o controle sobre os espaços urbanos se desloca para o campo digital, onde o acesso é regulado por sistemas de dados e redes de informação, em vez de por portões com guardas.

A instantaneidade da ubiquidade conduz à atopia de uma interface singular e após as distâncias de espaço e tempo, a distância de velocidade dissolve a noção de dimensão física (Virilio, 1993). Isso quer dizer que, a atenção se desloca do “onde” algo se encontra para uma realidade em que o essencial é a “velocidade” com que algo pode ser transmitido ou acessado. O “espaço” se torna menos relevante que a “velocidade de acesso”. Desse modo, vive-se em um mundo que se revela como uma obra do engenho humano, sendo mais artificial do que natural (Simon, 1996). Logo, os objetos, espaços urbanos, sistemas de transporte, redes digitais e até mesmo

os relacionamentos carregam as marcas desse artifício, sendo o resultado de um planejamento e manipulação humana.

Identificam-se, com essa conjuntura, quatro indícios que distinguem o artificial do natural, permitindo, dessa forma, delimitar os contornos das ciências do artificial (Simon, 1996). As criações artificiais são concebidas e sintetizadas pelos seres humanos, ainda que nem sempre de forma plenamente intencional ou premeditada (Simon, 1996). O artificial é feito pela mão humana, mas pode surgir tanto de projetos cuidadosos quanto de intervenções menos planejadas.

E as criações artificiais podem replicar a aparência dos elementos naturais, mas carecem, em um ou mais aspectos, da realidade destas últimas (Simon, 1996). A semelhança é superficial porque o artificial não compartilha plenamente da essência ou das propriedades fundamentais dos objetos naturais. Ademais, as atividades artificiais podem ser definidas com base em suas funções, objetivos e capacidade de adaptação (Simon, 1996). Demonstra-se que valor e a identidade das coisas artificiais estão diretamente ligados ao propósito para o qual foram criadas, sua função e a capacidade de se adaptarem a esse fim específico.

Inclusive, cita-se como exemplo o telefone celular, o qual foi criado inicialmente com o fito de facilitar a comunicação, mas suas características evoluíram de forma disruptiva. Hodiernamente, ele consiste em uma ferramenta multifuncional, visto que possuem câmeras avançadas, softwares de inteligência artificial e outras funções adicionais para atender os múltiplos propósitos advindos das demandas contemporâneas.

Ao idealizar algo artificial, considera-se tanto a descrição do objeto (como ele é) quanto às instruções e normas sobre seu funcionamento ideal (como ele deve se comportar ou o que deve realizar). Isso reflete a ideia de que coisas artificiais são frequentemente debatidas, especialmente durante sua concepção, tanto sob uma perspectiva normativa quanto descritiva (Simon, 1996). A realização do propósito ou a adaptação a um objetivo estabelece uma relação entre três elementos: o propósito ou objetivo, o caráter do artefato e o ambiente em que o artefato opera (Simon, 1996).

Nessa perspectiva, diversos profissionais das ciências biológicas, como fisiologistas, psicólogos e sociólogos, demonstram interesse pela cibernética e desejam aplicar seus métodos e técnicas em suas respectivas áreas de especialização (Ashby, 1956). Contudo, diversos foram desencorajados a investigar o tema, sob a crença de que seu entendimento exigiria uma profunda imersão em eletrônica e matemática avançada, nutrindo a

percepção de que a cibernética seria inextricavelmente vinculada a esses domínios (Ashby, 1956).

Isso remete que diversos profissionais de distintas áreas do saber possuíam interesse em aprender o funcionamento desse campo em virtude de suas possíveis contribuições promissoras. Pela percepção equivocada da exigência de haver um extenso conhecimento computacional e o domínio de cálculos complexos levou à ideia de que seria algo inacessível para aqueles sem formação técnica específica. Desse modo, o potencial de uma ferramenta interdisciplinar capaz de enriquecer várias áreas acabou, paradoxalmente, criando barreiras para sua ampla aplicação e disseminação.

Nos dias que correm, tornou-se ainda mais crucial que os juristas se envolvam no diálogo com a ciência da robótica com o intuito de compreender plenamente o funcionamento dos robôs e a extensão de sua autonomia em relação aos seus criadores e/ou aos que os utilizam em suas atividades profissionais (Monteiro, 2020). A interdisciplinaridade se torna não apenas desejável, mas imprescindível. Com essa circunstância, a transformação tecnológica no direito processual incita a nova geração de processualistas a cultivar a habilidade de se tornarem verdadeiros arquitetos dos tribunais virtuais (Nunes; Paolinelli, 2021). Aos juristas do presente e do porvir desenvolver competências digitais se torna uma obrigação inafastável. Caso contrário, a decisão da máquina será entendida unicamente por matemáticos e cientistas da computação, sem possibilidade de contestação (O'neil, 2020).

Os princípios fundamentais da cibernética podem ser abordados sem intervenção eletrônica e são essencialmente simples; destarte, embora técnicas sofisticadas sejam oportunas para aplicações avançadas, é possível grandes progressos - especialmente nas ciências realizadas biológicas - com métodos relativamente simples, desde que se possua uma compreensão clara e profunda dos princípios envolvidos (Ashby, 1956). Percebe-se que Ashby acredita na potencialidade da cibernética como uma disciplina acessível e aplicável de maneira ampla mesmo sem dependência de conhecimentos labirínticos.

Para ele, o ponto crucial para essa aprendizagem é a compreensão sólida de seus princípios fundamentais, talvez seria como seguir o “coelho branco” da obra literária “Alice no País das Maravilhas” (Carroll, 1865), que também é referência simbólica no filme “*Matrix*”, onde Neo, o protagonista (engenheiro de *software* por profissão e *hacker* por interesse

pessoal), é convidado a seguir o animal para começar a trajetória de descobertas de uma realidade oculta.

A circunstância envolve a decisão de sair da zona de conforto para adentrar em um processo de conhecimento do território misterioso que desafia pré-compreensões. É ir além do superficial em preceitos basilares que possibilitam o entendimento de suas essências. Posto que, a verdadeira essência e a correta compreensão da profundidade do logos preservam a possibilidade de que ele se amplie por si mesmo (Streck, 2014). É justamente esse movimento que se propõe: investigar, com base sólida e crítica, a complexa interseção entre direito e inteligência artificial. O futuro exige uma nova disciplina jurídica transversal, fluida e comprometida com os princípios que regem o novo paradigma.

Nesse contexto, escolhe-se embarcar em uma jornada rumo ao desconhecido em função da complexidade crescente que envolve a interseção entre a inteligência artificial e o direito. A articulação entre diferentes áreas do saber exige, de forma cada vez mais evidente, a constituição de uma disciplina jurídica própria, de natureza multidisciplinar/interdisciplinar, fundamentada nos princípios que regem a inteligência artificial (Peixoto; Silva, 2019).

2.1 Navegando com a Nabucodonosor pelos algoritmos: investigando as dimensões do sistema de comportamento inteligente

Dá-se início a uma viagem rumo às profundezas dos sistemas de comportamento inteligente, com o fito de desconstruir a noção simplificada do termo “algoritmo”, frequentemente tratado como um elemento singular. Nesta jornada, busca-se compreender não apenas os pilares que sustentam esses sistemas, mas também o funcionamento intrincado de suas múltiplas camadas. Esses elementos, longe de serem isolados, operam como engrenagens que, em sequência coordenada, tornam possível a execução de tarefas complexas e a tomada de decisões automatizadas. Explorar essa arquitetura é essencial para compreender como os sistemas de inteligência artificial se estruturam e impactam o campo jurídico.

Nessa senda, um programa de computador é fruto da atividade intelectual do programador e todo programa de computador tem origem em um algoritmo, isto é, na organização lógica de instruções destinadas à resolução do problema (Filho; Lindemann, 2013). A programação é um

processo intelectual e técnico, que exige conhecimento e planejamento para que exista uma estrutura eficiente. Desse modo, um algoritmo consiste em uma sequência finita de etapas destinadas à solução de um problema específico (Filho; Lindemann, 2013). Cada algoritmo é uma tentativa de decompor uma circunstância em partes menores e mais simples, sendo o esqueleto de um programa.

Essa lógica está presente não apenas no campo tecnológico, mas também em reflexões filosóficas, dado que na obra *Ética a Nicômaco* (Livro III.3, 1112b), Aristóteles aprofunda esse tema ao sugerir um algoritmo (Russell; Norvig, 2013). Embora o termo “algoritmo” não existisse à época, o filósofo delineia uma espécie de processo racional sequencial. Essa concepção é ilustrada quando é afirmado que “não deliberamos acerca de fins, mas a respeito de meios” (Aristóteles, 1991, p.53). Posto que um médico, por ocasião, não delibera sobre curar ou não, um orador sobre persuadir, um estadista sobre implantar a ordem pública, assim como nenhum outro delibera a respeito de sua nomeação (Aristóteles, 1991). Então, sugere-se que os profissionais não devem questionar se o resultado final é desejado, mas sim como alcançá-lo de maneira adequada. A escolha dos meios é tão importante quanto o objetivo a ser alcançado.

Costuma-se resolver algoritmos diariamente, de maneira tão automatizada que nem se percebe. Qualquer atividade do dia a dia que envolva uma sequência de passos para sua realização pode ser considerada um algoritmo (Filho; Lindemann, 2013). Por exemplo, substituir uma lâmpada, preparar um bolo, fazer um café coado, escovar os dentes, trocar um pneu, entre outros. (Filho; Lindemann, 2013). Permite-se, assim, observar essa estrutura de raciocínio como um elemento essencial no processo de tomada de decisões e na resolução de problemas cotidianos. De fato, parece que a própria mente humana segue um conjunto de “algoritmo mental” para lidar com as tarefas diárias. O fato da execução dessas atividades de forma automática revela o quão profundamente são internalizadas essas sequências, a ponto de não mais as reconhecer como processos complexos, mas como gestos naturais e quase instintivos.

Para elaborar um algoritmo, é essencial compreender claramente o problema a ser resolvido, ou seja, definir com precisão o objetivo que se pretende alcançar e para isso é preciso verificar o problema identificando três elementos basilares: entrada (*Input*), processamento e saída (*Output*) (Filho; Lindemann, 2013). Há uma lógica sequencial que organiza as etapas do raciocínio. A entrada corresponde aos elementos fundamentais para o

início da resolução do problema, o processamento envolve a cuidadosa manipulação desses elementos a fim de obter a saída, que se apresenta como o resultado final do processo (Filho; Lindemann, 2013).

Nesse sentido, considerando que o direito processual civil, tal como é reconhecido na contemporaneidade, encontra uma de suas fontes no direito praticado na Roma Antiga, sobretudo no período justiniano (século VI d. C.) (Isaia, 2012). A gênese do rito ordinário nasceu nesse tempo e estabeleceu uma distinção entre julgamento e decisão, onde é separando claramente a fase de julgamento, na qual ocorre o estudo dos fatos e argumentos das partes, da etapa decisória, na qual o magistrado emite sua sentença fundamentada.

A ilustração supramencionada seria como um andamento processual em que a petição e os documentos anexados representam a entrada, fornecendo os elementos necessários para iniciar um processo judicial. O processamento ocorre durante o trâmite processual, abrangendo aspectos como o estudo de documentos, audiências, produção de provas e demais etapas. Por fim, a saída equivale à sentença que se constitui como a resposta jurídica para a situação posta.

Ao seguir o itinerário, em 1936, Alan Mathison Turing desenvolveu o estudo de uma máquina universal para descrever formalmente os conceitos de computação e algoritmos. Nesse estudo Turing afirma que a memória humana é, por sua natureza, intrinsecamente limitada, mas é possível inventar uma única máquina capaz de calcular qualquer sequência computável (Turing, 1936). Assim, o teste de Turing, idealizado por Alan Turing (1950), foi concebido para oferecer uma definição operacional satisfatória de inteligência (Russell; Norvig, 2013).

O computador será considerado aprovado no teste se um interrogador humano, ao formular perguntas por escrito, não for capaz de discernir se as respostas recebidas são originadas de uma pessoa ou de uma máquina (Russell; Norvig, 2013). Assim, o teste consiste em fazer com que um programa conduza uma conversação, por meio de mensagens digitadas on-line, com um interrogador durante cinco minutos e o interrogador deve, então, discernir se estava conversando com um programa ou uma pessoa; o programa é considerado aprovado no teste se conseguir enganar o interrogador por, no mínimo, 30% do tempo (Russell; Norvig, 2013). O foco não está apenas na compreensão ou geração de respostas, mas na criação de uma ilusão de entendimento, o que levanta questões sobre a própria natureza da inteligência e da percepção humana.

Por ora, constata-se que programar um computador para obter êxito no teste demanda considerável esforço (Russell; Norvig, 2013). Em virtude de que a mera imitação superficial da linguagem não é suficiente para persuadir um interlocutor atento. Para que o computador atenda às exigências, seria necessário que possuísse algumas capacidades fundamentais como: o processamento de linguagem natural (PLN) que seria essencial para permitir que ele se comunicasse de maneira eficiente em um idioma natural; a representação de conhecimento para o armazenamento das informações que sabe ou ouve; o raciocínio automatizado, o qual capacita a utilização das informações armazenadas para responder a perguntas e formular novas instruções; e o aprendizado da máquina para adaptação às novas situações e que consiga identificar e extrapolar padrões (Russell; Norvig, 2013). Sem esse mecanismo, qualquer progresso seria limitado a um conjunto finito de respostas, tornando evidente a artificialidade da interação.

Todavia, o chamado teste de Turing total incorpora um sinal de vídeo permitindo que o interrogador avalie as habilidades de percepção do indivíduo além de oferecer a possibilidade de interagir com objetos físicos simulando a entrega desses itens “pela janelinha” (Russell; Norvig, 2013). O que desloca o paradigma da inteligência artificial para além da cognição linguística, adentrando o domínio da corporeidade e da sensopercepção. E para ser aprovado no teste de Turing total, o computador deverá dispor de visão computacional para refletir objetos e de robótica para manipulá-los e se movimentar no ambiente (Russell; Norvig, 2013).

No final da década de 1940 surgiu a cibernética. Norbert Wiener define a cibernética como “a ciência do controle e da comunicação, no animal e na máquina”, resumindo-a em uma expressão: a arte da pilotagem (Ashby, 1956). Remete-se à ideia de conduzir e controlar um sistema complexo de forma eficaz, como ilustrado na obra clássica “A Arte da Guerra”, de Sun Tzu, onde a estratégia, o domínio do ambiente e a adaptação constante são essenciais para o sucesso no comando. Wiener ao definir Cibernética no livro original, incluiu comunicação e controle na mesma categoria, atentando que quando ocorre uma comunicação entre duas pessoas, uma transmite uma mensagem para outra que, ao responder, envia uma mensagem conectada contendo informações que lhe são originalmente acessíveis, mas não ao interlocutor inicial (Wiener, 1954).

Para elucidar, de forma exemplificativa, esse processo ocorre quando Neo pergunta para Morpheus, seu mentor, “O que é Matrix?” e ele obtém

a resposta “É Controle”. Esse diálogo ocorre após um ponto crítico e épico de decisão que é a escolha entre a pílula vermelha (que o levaria à verdade, simbolizando luta e agitação) e a pílula azul (associada à tranquilidade e permitiria continuar vivendo na ilusão). Ele opta pela primeira, iniciando sua expedição em busca da verdade e marca sua ruptura definitiva com a realidade virtual que até então acreditava ser o mundo real.

A compreensão da sociedade só é possível por meio do estudo das mensagens e dos recursos de comunicação disponíveis e no desenvolvimento futuro dessas mensagens e meios de comunicação como interações entre homem e máquina entre máquina e homem e entre máquinas estão projetadas para assumir um papel cada vez mais relevante (Wiener, 1954). O que torna evidente a necessidade de embarque na Nabucodonosor, a nave de Morpheus, para começar a assimilação do papel das máquinas e o impacto de suas interações no mundo.

Na parábola da Caverna de Platão ocorre uma situação semelhante, quando homens vivem no ambiente subterrâneo, voltados de costas para a entrada e acorrentados de forma que não conseguem se mover. Tudo o que podem ver é a parede à sua frente. Atrás deles, há um muro que os separa da entrada e do outro lado há figuras que passam carregando objetos, projetando sombras na parede. Então, a única percepção possível aos homens aprisionados é esse “teatro de sombras” que se desenrola diante de seus olhos. E eles estão ali desde o nascimento e as sombras representam tudo o que existe (Gaarder, 1995).

Isso representa como, muitas vezes, os seres humanos se restringem a uma visão limitada do mundo, aceitando como verdade única aquilo que percebem diretamente, sem questionar ou explorar além das aparências. Contudo, supõe-se que um dos prisioneiros da caverna tenha conseguido se libertar de sua prisão. Primeiramente, questiona-se a origem dessas imagens projetadas na parede da caverna e ofuscado pela luz brilhante (o sol), ele esfregaria os olhos e, então, perceberia a beleza de tudo, vendo pela primeira vez cores e contornos nítidos (Gaarder, 1995).

Ao pensar em todos aqueles que ainda estão na caverna ele retorna e, ao chegar, tenta explicar aos outros prisioneiros que as sombras na parede são apenas cópias trêmulas de coisas verdadeiras, mas ninguém acredita nele (Gaarder, 1995). Sua tentativa de mostrar que há algo além das sombras causa inquietação entre os prisioneiros e incompreendido acaba sendo morto. Isso pode ser visto como uma crítica à sociedade que, ao temer a verdade e as mudanças que ela pode trazer, tenta silenciar os

que buscam iluminar novos caminhos. Pensa também em Sócrates, que os “habitantes da caverna” assassinaram, eliminando aquele que poderia revelar o verdadeiro caminho do conhecimento, ao desafiar as crenças convencionais (Gaarder, 1995). Na medida que sob tal mito, descortina-se a busca pelo conhecimento e pela verdade, ilustrando a dificuldade de questionar e superar preconceitos e convenções que limitam a percepção.

Na transição para a era contemporânea, os algoritmos ergueram uma “parede” moderna que, tal como na narrativa supracitada, confina indivíduos em bolhas de conteúdos filtrados e moldados segundo perspectivas específicas. Essa filtragem constante reforça vieses, perpetuam crenças e dificulta a expansão para um horizonte mais amplo de conhecimento e compreensão crítica da realidade. Por isso, torna-se necessária a busca por um esforço consciente para “sair da caverna” digital, questionando as “verdades” apresentadas e promovendo uma visão mais abrangente e crítica.

A Inteligência Artificial (IA) é um campo da ciência computacional que se dedica à automação do comportamento inteligente (Luger, 2013). Melhor explicando, esse campo de estudo visa desenvolver sistemas e algoritmos capazes de simular funções cognitivas humanas, como aprendizado, racionalidade e tomada de decisões. Dessa forma, a IA é definida como o estudo de agentes que recebem percepções do ambiente e realizam ações (Russell; Norvig, 2013). Trata-se, portanto, de uma capacidade operacional e adaptativa frente aos estímulos do meio de inserção.

Presume-se que o programa será executado em um dispositivo de computação dotado de sensores e atuadores físicos, ao que se refere como uma arquitetura: *Agente = arquitetura + programa* (Russell; Norvig, 2013). Um agente é, essencialmente, algo que age (a palavra “agente” deriva do latim *agere*, que significa “fazer”) e é tudo aquilo que é capaz de perceber seu ambiente por meio de sensores e de interagir com ele por meio de atuadores (Russell; Norvig, 2013). Em termos mais simples, a arquitetura se refere ao dispositivo de computação com sensores e atuadores que permitem a interação física com o ambiente, enquanto o programa é a parte lógica que determina como o agente deve reagir a essas percepções.

Por sinal, quando um agente depende do conhecimento prévio de seu projetista, em vez de se basear em suas próprias percepções, afirma-se que o agente carece de autonomia (Russell; Norvig, 2013). Ao depender do conhecimento de seu projetista, o agente artificial não está lançado no

mundo de forma autêntica e seu horizonte de compreensão é derivado e fechado. De modo geral, a arquitetura disponibiliza ao programa as percepções captadas pelos sensores, executa o programa e transmite as ações selecionadas pelo programa aos atuadores conforme elas são geradas (Russell; Norvig, 2013). A autonomia não se resume apenas à ideia de “independência”, mas à capacidade de adaptação e aprendizado contínuo.

Nessa conjuntura, uma sociedade hierárquica pode ser comparada a uma árvore, onde o agente de cada ramo assume a responsabilidade exclusiva pelos agentes que dele se derivam (Minsky, 1986). Existe um processo de delegação e responsabilidade dentro de uma rede de interações. Esse padrão é amplamente observado em diversas áreas, pois segmentar o trabalho dessa forma constitui, frequentemente, a abordagem mais eficaz para iniciar a resolução de um problema (Minsky, 1986). A hierarquia simplifica a distribuição de responsabilidades e a tomada de decisões, proporcionando uma maneira ordenada de enfrentar complexidades que, de outro modo, seriam de difícil manejo. Os algoritmos destinados à construção de árvores de decisão figuram entre os mais reconhecidos e amplamente utilizados (Salzberg, 1994).

Quinlan aborda uma variedade de questões relacionadas às árvores de decisão, abrangendo desde o algoritmo principal para a construção da árvore inicial até técnicas de poda, conversão de árvores em regras e o tratamento de desafios como a ausência de valores nos atributos (Salzberg, 1994). Existem técnicas de refinamento e otimização, como a supressão que visa remover ramos desnecessários ou excessivos, aprimorando a generalização da árvore e evitando o sobreajuste (*overfitting*).

Algoritmos de árvore de decisão iniciam com um conjunto de casos ou exemplos e constroem uma estrutura de árvore de dados, a qual pode ser empregada para classificar novos casos e cada caso é caracterizado por um conjunto de atributos (recursos), seja numéricos ou simbólicos, e vinculado a um rótulo que designa o nome de uma classe (Salzberg, 1994). A partir desse conjunto inicial, o algoritmo constrói uma estrutura hierárquica, representada como uma árvore que pode ser utilizada para classificar novos casos com base no conhecimento adquirido durante o treinamento.

Cada nó interno de uma árvore de decisão apresenta um teste, cujo resultado determina o ramo a ser seguido a partir desse nó (Salzberg, 1994). O critério de escolha pode variar, mas geralmente é baseado em medidas de impureza, como a *Gini Index* ou a *Entropia*. Esses critérios avaliam o grau

de homogeneidade dos dados após a divisão, preferindo sempre as divisões que resultam em grupos mais homogêneos, ou seja, que são mais puros em relação à classe final. Então, é possível acompanhar o caminho da decisão através da árvore e entender a previsão verificando as contribuições de cada nó (Nunes; Andrade, 2023).

Embora seja considerado que o campo da inteligência artificial (IA) teve como início oficial em meados da década de 1950, as questões intelectuais relevantes remontam a períodos bem anteriores (Newell, 1983). Os debates fundantes da IA atravessam séculos e perpassam campos como a filosofia, a lógica, a psicologia, a neurofisiologia e a ciência da computação. Trata-se de um percurso histórico rico em dissensos e rupturas, nos quais conceitos como emoção, linguagem, aprendizagem, símbolos e algoritmos foram tensionados e reconfigurados, conforme o quadro abaixo que sistematiza os principais marcos dessas controvérsias.

Quadro 01 - Panorama histórico das principais questões intelectuais no campo da Inteligência Artificial.

ANO	FATO
1640-1945	Mecanismo versus teleologia: resolvido com cibernética
1800-1920	Biologia natural versus vitalismo: estabelece o corpo como uma máquina 1870- Razão versus emoção e sentimento (separa as máquinas dos homens).
1870-1910	Filosofia versus ciência da mente: separa a psicologia da filosofia.
1910-1945	Lógica versus psicológica: separa a lógica da psicologia.
1940-1970	Analógico versus digital: cria a ciência da computação.
1955-1965	Símbolos versus números: isola IA na ciência da computação.
1955	Sistemas simbólicos versus contínuos: separa a IA da cibernética
1955-1965	Resolução de problemas versus reconhecimento: separa a IA do reconhecimento de padrões.
1955-1965	Psicologia versus neurofisiologia: separa a IA da cibernética.
1955-1965	Desempenho versus aprendizagem: divide a IA do reconhecimento de padrões
1955-1965	Serial versus parallel: coordene com os quatro problemas acima.
1955-1965	Heurísticas versus algoritmos: isola IA na ciência da computação.
1955-1985	Interpretação versus compilação: isola IA na ciência da computação.

ANO	FATO
1955	Análise de simulação versus engenharia: divide a IA 1960-Substituindo versus ajudando humanos: isola a IA.
1960	Epistemologia versus heurística: divide a IA (menor); conecta-se com a filosofia.
1965-1980	Pesquisa versus conhecimento: aparente mudança de paradigma dentro da IA.
1965-1975	Poder versus generalidade: mudança de tarefas de interesse.
1965	Competência versus desempenho: separa a linguística da IA e da psicologia.
1965-1975	Memória versus processamento: psicologia cognitiva dividida da IA.
1965-1975	Resolução de problemas versus reconhecimento: reconhecimento se junta à IA via robótica.
1965-1975	Sintaxe versus semântica: separa a linguística da IA.
1965	Prova de teoremas versus resolução de problemas: divide IA.
1965	Engenharia versus ciência: divide a ciência da computação, incluindo IA.
1970-1980	Linguagem versus tarefas: a linguagem natural torna-se central.
1970-1980	Representação processual versus declarativa no 1: mudança da prova de teoremas.
1970-1980	Quadros versus átomos: mudança para representações holísticas.
1970	Razão versus emoção e sentimento: separa a IA da filosofia da mente.
1975	Brinquedo versus tarefas reais: mudança para aplicativos.
1975	Serial versus parallel: IA distribuída (sistemas semelhantes a boatos).
1975	Desempenho versus aprendizagem: ressurgimento (sistemas de produção)
1975	Psicologia versus neurociência: newlinkto neuroscience.
1980	Serial versus parallel: nova tentativa de sistemas neurais;
1980	Resolução de problemas versus reconhecimento: retorno da robótica;
1980	Representação processual versus declarativa no 1: PROLOG;

Fonte: Newell (1983)

Os 27 períodos destacados refletem marcos significativos no desenvolvimento das ciências, especialmente no campo da inteligência

artificial e das ciências cognitivas. Sublinha-se, especialmente alguns, como quando ocorreu a separação entre uma diferenciação entre a razão em relação à emoção e o sentimento. Atribui-se razão (lógica fria) às máquinas e se acredita que uma máquina não poderia ter coração, nem sentimentos ou emoções que pudessem conflitar com sua razão é o que causa preocupação (Newell, 1983).

Os seres humanos artificiais do século XIX pareciam reunir todas as características humanas, como exemplificado pelo monstro criado por Frankenstein, e no século XX, especialmente em R.U.R. (Robôs Universais de Rossum), encontra-se claramente o robô inteligente, que não possui alma e, portanto, carece de emoções ou desejos próprios (Newell, 1983). O monstro de Frankenstein, por sua própria natureza, é uma representação simbólica da tentativa de criar um “ser humano”, mas falha ao não possuir as qualidades que definem a humanidade, como a alma e a consciência. Este exemplo enfatiza a separação entre a inteligência mecânica e a humanidade devido às limitações das máquinas em relação aos seres humanos. A relevância disso para a IA reside em oferecer uma base para distinguir as máquinas dos humanos, o que se difere da questão do propósito (Newell, 1983).

Entre 1870 e 1910, ocorreu a distinção entre psicologia e filosofia. “O que restou foi uma tradição filosófica contínua, focada na mente, que se mostrou completamente distinta do campo da psicologia e, ainda mais, da tecnologia” (Newell, 1983, p. 194). A psicologia focou em processos mentais de uma forma empírica, enquanto a filosofia permaneceu mais voltada para questões de consciência e subjetividade. Essa separação garantiu que, quando a IA surgisse, o que aconteceu instantaneamente quando os computadores se tornaram suficientemente poderosos, sua relação com a filosofia da mente seria apenas periférica (Newell, 1983).

Quando os computadores foram inicialmente concebidos na década de 1940, dividiram-se em duas grandes categorias: os computadores analógicos representavam quantidades por meio de variáveis físicas contínuas, como corrente ou tensão; destacavam-se pela velocidade e processamento simultâneo, embora possuísem precisão inerentemente limitada (Newell, 1983). Já, os computadores digitais codificavam informações em estados discretos, operando de forma sequencial e, apesar de mais lentos, apresentavam precisão teoricamente ilimitada” (Newell, 1983). O advento da computação moderna marcou o início de uma divisão estrutural desses sistemas, motivada pela necessidade de explorar

distintas abordagens na resolução de problemas matemáticos e científicos com maior eficiência.

Com o tempo, surgiram os computadores híbridos, concebidos para reunir as vantagens de ambos os sistemas: a precisão e o controle digital aliados à velocidade e praticidade dos dispositivos analógicos, porém a partir da década de 1970, o termo “computador” passou a referir-se exclusivamente às máquinas digitais (Newell, 1983). O fim da era analógica não foi determinado apenas pelo aumento da velocidade e da eficiência de custos dos sistemas digitais, mas também pela descoberta da Transformada Rápida de Fourier, que fundou o campo do processamento de sinais digitais e, assim, penetrou no mais sólido reduto da computação analógica (Newell, 1983). Esse estilo de interpretar pertence ao mundo da cibernética e não ao da IA (Newell, 1983). Dessa forma, a transição da computação analógica para a digital está mais associada ao universo da cibernética, pois envolve a evolução dos mecanismos computacionais e das arquiteturas de processamento e não necessariamente a criação de sistemas capazes de replicar o pensamento humano.

Nesse percurso, quando ocorre a distinção entre símbolos e números, chega-se agora à primeira das questões que caracterizam a própria IA, em oposição ao pano de fundo contra o qual surgiu (Newell, 1983). O domínio da computação digital consolida a definição de computadores como máquinas destinadas à manipulação de números e seus defensores ressaltaram que a grande virtude desse modelo residia na possibilidade de codificar qualquer tipo de informação em formato numérico, inclusive as próprias instruções de operação (Newell, 1983). Isso remete que ao traduzir as instruções e as informações para um formato numérico, as máquinas podem ser programadas para executar uma variedade de tarefas.

Em contraposição, os cientistas da IA consideravam os computadores como máquinas capazes de manipular símbolos e a grande vantagem, segundo eles era que tudo podia ser codificado em símbolos, até os mesmos números (Newell, 1983). Os pesquisadores em IA se orgulhavam do fato de que seus programas, apesar de sua complexidade, não necessitavam de multiplicações, sendo suficientemente sofisticados para provar teoremas ou jogar jogos (Newell, 1983). O que indicava uma forma mais próxima do raciocínio humano, onde ideias e conceitos são frequentemente expressos de forma simbólica e não apenas numérica (refletindo uma tentativa de imitar, de certa forma, a cognição humana).

A grande parte do campo da informática, assim como seus principais líderes de pensamento, aceitou sem questionamentos a concepção de que os computadores são meras máquinas que manipulam números e os investigadores em IA, por sua vez, estavam profundamente empenhados em promover uma nova perspectiva, considerando a visão tradicional como uma interpretação radicalmente equivocada da natureza do computador e uma barreira substancial à ideia de que as máquinas poderiam ser inteligentes (Newell, 1983). Essa divisão, por sua vez, teve implicações significativas na forma como a IA evoluiu, influenciando tanto sua trajetória de pesquisa quanto a forma como ela está integrada (ou não) a outras áreas do conhecimento.

De forma sucessiva, uma dicotomia entre sistemas simbólicos e contínuos se entrelaça com a oposição entre psicologia e neurofisiologia para separar a IA da cibernética. No primeiro fato, na década de 1960, o principal efeito histórico dessa questão foi a separação quase total entre aqueles que concebiam sistemas como contínuos e aqueles que se concentravam em sistemas de programação (Newell, 1983). O pessoal do sistema contínuo foi alocado em departamentos de engenharia elétrica enquanto os pesquisadores de IA se estabeleceram em departamentos da ciência da computação (Newell, 1983).

Então, os que trabalhavam com a abordagem de sistemas contínuos focavam no reconhecimento de padrões como a principal tarefa a ser desempenhada por máquinas, incluindo o reconhecimento de caracteres, fala e padrões visuais e quase sempre, o conhecimento adquirido estava relacionado com uma capacidade de reconhecimento que estava sendo aprendida (Newell, 1983). Reforça-se a ideia de que o aprendizado ocorre por meio da identificação e processamento de dados sensoriais.

Ao contrário, aqueles que operavam dentro do paradigma dos sistemas simbólicos se concentravam na resolução de problemas, encarando-a como uma tarefa primordial para que as máquinas jogassem, provassem teoremas e solucionassem quebra-cabeças. (Newell, 1983). Utiliza-se de raciocínio formal e estruturado para entender e resolver diferentes desafios, em vez de simplesmente considerar padrões no ambiente. Por exemplo, em uma partida de xadrez é necessário seguir regras para vencer como no *puzzle* é necessário combinar peças de forma lógica.

Na psicologia e na neurologia (como é bem sabido, essas duas disciplinas se expressam de maneiras completamente distintas, embora não necessariamente em oposição) os estudiosos de sistemas contínuos e

reconhecimento de padrões retornaram para a neurofisiologia, enquanto aqueles dedicados aos sistemas simbólicos e à resolução de problemas, ou seja, à IA – buscaram respaldo na psicologia (Newell, 1983). Nessa maré, a preferência da IA pela psicologia decorreu do fato de que os sistemas de programação com os quais ela lidava envolviam tarefas de natureza essencialmente psicológicas, e não fisiológicas (Newell, 1983). Assim, enquanto a neurociência buscava uma compreensão do “como” fisiológico da mente, a psicologia oferece *insights* sobre o “como” funcional e “porquê” das decisões.

Mais adiante os dissensos entre heurísticas e algoritmos, bem como entre interpretação e compilação contribuíram para isolar a IA da ciência da computação. Os algoritmos passaram a ser concebidos como programas capazes de resolver um problema ou solucioná-lo dentro de limites de tempo predefinido e os programas eficazes demonstravam uma estrutura algorítmica robusta, caso contrário, a responsabilidade recaía sobre o programador, que não havia estudado o problema com a profundidade necessária para definir o que o programa deveria executar para solucioná-lo (Newell, 1983).

Em contrapartida, os programas heurísticos operavam com base em regras práticas de caráter heurístico e a abordagem heurística denota uma escolha por um caminho distinto daquele que busca a solução ideal por meio da engenharia ou da análise matemática de um problema (Newell, 1983). Além disso, fornecer a um programa apenas algumas regras parciais e não completamente analisadas era considerado, na melhor das hipóteses, estranho, e, na pior, irracional ou até frívolo (Newell, 1983). A incansável busca pela perfeição técnica, em contraste com a necessidade pragmática de solucionar problemas reais, evidencia que as limitações computacionais tornam inviáveis soluções plenamente otimizadas.

Os programas de IA foram criados em linguagens de processamento de listas, que operavam de forma interpretativa, enquanto a principal tendência no desenvolvimento das linguagens se direcionava de maneira irreversível para o uso de compiladores (Newell, 1983). Antes de meados da década de 1950, as linguagens de programação além dos montadores eram interpretativas e o grande marco na evolução dos compiladores ocorreu com o desenvolvimento do Fortran (*Formula Translation*), na metade dos anos 1950, definindo os rumores do avanço das linguagens de programação (Newell, 1983). Diferentemente das linguagens interpretativas, o compilador converteu previamente o código-fonte em

um formato incluído, proporcionando ganhos expressivos em desempenho e eficiência. A propósito, no início de abril de 2021, a linguagem de programação Fortran, desenvolvida pela IBM, alcançou a 20ª posição no ranking das linguagens mais populares, conforme o índice da Comunidade de Programação Tiobe, que monitora a relevância e o uso dessas linguagens (UFSM, 2021).

Nesse contexto, a velocidade de execução era a principal prioridade para a comunidade de programadores, em contrapartida, a inteligência artificial valorizava o caráter interpretativo de suas linguagens, considerando-o essencial para o desenvolvimento de sistemas inteligentes (Newell, 1983). Essa abordagem se manifestou na adoção da inteligência completa e influenciou toda a filosofia de design das linguagens, sustentando a ideia instigante de incorporar o próprio intérprete (Newell, 1983). Isso se deve ao fato de que a manipulação simbólica, central na IA, exige um ambiente computacional dinâmico e adaptável, algo que linguagens interpretativas oferecem com mais facilidade do que aquelas focadas exclusivamente na eficiência computacional. Essa dissociação entre a programação em inteligência artificial e as linguagens de alto nível convencionais, instaurada desde o surgimento da IA, mantém-se vigente até os dias atuais, gerando repercussões que se estendem de maneira ainda mais profunda do que se poderia inicialmente prever (Newell, 1983).

Curiosamente, os esforços nessa área tiveram início logo após o término da Segunda Guerra Mundial, e a própria designação “Inteligência Artificial” foi introduzida em 1956 no Seminário de Pesquisa de Verão de Dartmouth sobre Inteligência Artificial, John McCarthy convenceu Marvin Minsky, Claude Shannon e Nathaniel Rochester a se juntarem a ele na missão de reunir pesquisadores norte-americanos específicos na teoria dos autômatos, redes neurais e no estudo da inteligência (Russell; Norvig, 2013). O objetivo do estudo foi a partir da premissa fundamental de que qualquer aspecto da aprendizagem ou qualquer outra característica da inteligência possa ser descrito com tal precisão que seja possível construir uma máquina capaz de simulá-lo (Russell; Norvig, 2013). A partir desse encontro seminal é possível perceber que muitas das questões levantadas naquela época permanecem relevantes.

Houve uma tentativa de compreender como capacitar as máquinas para utilizar a linguagem, a partir de abstrações e conceitos, para resolver os tipos de problemas que atualmente são exclusivos dos seres humanos e para se aperfeiçoar, mas uma dupla de pesquisadores da Carnegie Tech, Allen

Newell e Herbert Simon, simplesmente roubaram a cena (Russell; Norvig, 2013). Apesar de outros pesquisadores terem idealizado e, em alguns casos, desenvolvido programas voltados para aplicações específicas, como jogos de damas, Newell e Simon avançaram significativamente ao criar o *Logic Theorist* (LT), um programa capaz de raciocinar e sobre essa realização, Simon destacou: “Desenvolveu-se um programa de computador que pode pensar de forma não numérica, resolvendo, assim, o antigo dilema mente-corpo.” (Russell; Norvig, 2013). Essa declaração não apenas enfatizava a importância do projeto, mas também sugeria que a barreira entre a cognição humana e a computação estava sendo desafiada.

Embora o Seminário de Dartmouth não tenha apresentado conceitos inéditos, seu maior legado foi conectar os principais protagonistas da história da Inteligência Artificial e nos vinte anos seguintes, esse grupo, ao lado de seus alunos e colaboradores do MIT, da CMU, de Stanford e da IBM, tornou-se a força motriz por trás do desenvolvimento da área, influenciando seus caminhos e avanços (Russell; Norvig, 2013). Ao construir essa alquimia do conhecimento, onde a convergência dos pontos de vista individuais se torna a semente para novas formas de pensar e interpretar o mundo, rompe com a ideia de que o conhecimento é uma construção unívoca e imutável.

Ao atentar para a proposta do seminário de Dartmouth se torna visível o motivo que era necessário que a IA se tornasse um campo separado (Russell; Norvig, 2013). Primeiramente, porque a IA desenvolveu desde o início a ambição de reproduzir faculdades humanas, como criatividade, autoaperfeiçoamento e uso da linguagem, áreas que não eram abordadas por outros e a segunda resposta é a metodologia (Russell; Norvig, 2013). A proposta central era a de que processos complexos da mente humana poderiam ser descritos com precisão suficiente para serem implementados em sistemas computacionais, abrindo caminho para a criação de máquinas que não apenas executassem tarefas, mas que aprendessem, se adaptassem e, de certa forma, “pensassem”.

Esse movimento inovador não se tratava apenas de uma evolução teórica, mas sim de uma reestruturação metodológica que propiciou o desenvolvimento de algoritmos e modelos que hoje sustentam os avanços da IA. A IA é o único desses campos que se estabelece de forma clara como um ramo da ciência da computação, embora a pesquisa operacional também enfatize simulações computacionais e a IA se distingue por ser o único campo de buscar a construção de máquinas capazes de operar de

maneira autônoma em ambientes complexos e dinâmicos (Russell; Norvig, 2013).

Alguns anos após o surgimento da Inteligência Artificial, emergiu um debate fundamental: a meta deveria ser a criação de sistemas capazes de substituir integralmente os humanos ou, ao contrário, o aprimoramento das interações entre pessoas e computadores? Embora a natureza ética dessa questão seja inegável, ela não foi inicialmente formulada como um dilema de ordem social, mas sim como uma escolha baseada em preferências individuais (Newell, 1983). Nessa bifurcação, a IA poderia ser considerada como um elemento que propicia uma parceria entre humanos e máquinas ou uma condição de domínio total das máquinas sobre as atividades humanas.

Com esse rumor, predispõe-se a necessidade de reflexão sobre os limites, os benefícios e as ameaças da dependência tecnológica crescente. No entanto, frequentemente se percebia um tom que, se não expressava uma superioridade ética explícita, ao menos refletia uma concordância com os ideais mais elevados da área e aqueles mais inclinados à Inteligência Artificial não enfrentaram essa questão de maneira direta, optando, em grande parte, por ignorá-la (Newell, 1983, p. 205). Sugerindo que os pesquisadores acreditavam estar conduzindo um projeto de grande relevância para a humanidade, motivado pelo desejo de expandir as capacidades intelectuais humanas ou mesmo de criar máquinas capazes de raciocinar. No entanto, essa convicção não significava que houvesse uma visão crítica sobre as consequências desse desenvolvimento.

Ao passo que a inteligência de máquina se tornou uma parte tão intrínseca do nosso cotidiano, quando se dá uma ordem a uma máquina, a situação não se distingue essencialmente daquela em que se dá uma ordem para uma pessoa (Wiener, 1954). Por outras palavras, na medida em que a consciência alcança, está ciente da ordem dada e do sinal de concordância recebido, para quem se refere, o fato de o sinal, em seus estágios intermediários, ter passado por uma máquina em vez de por uma pessoa é irrelevante e em nenhum momento altera significativamente a relação com o sinal (Wiener, 1954). A ideia central é que independentemente de a resposta passar por uma máquina ou por uma pessoa, isso não altera de forma significativa a relação entre quem dá a ordem e a resposta.

A inteligência artificial é comumente classificada em três categorias principais: inteligência artificial estreita ou fraca (ANI), inteligência artificial geral ou forte (AGI) e superinteligência artificial (ASI) (Webb,

2020). A propósito, os filósofos denominam como IA fraca a hipótese de que talvez as máquinas possam agir de maneira inteligente (ou até mesmo parecer inteligentes), enquanto a afirmação de que tais máquinas realmente pensam (em vez de simplesmente simular o pensamento) é referida como a hipótese de IA forte (Russell; Norvig, 2013). A IA fraca realiza tarefas mais específicas e é a mais presente diariamente como as assistentes virtuais Siri, Alexa e Google Assistant ou Software de Reconhecimento de Imagem e jogos de xadrez.

Inclusive, no confronto decisivo da sexta partida do embate ocorrido em 1997 entre o então campeão mundial de xadrez, Garry Kasparov, e o programa Deep Blue, Kasparov, conduzindo as peças pretas, viu-se obrigado a abandonar a partida, registrando, assim, a primeira ocasião em que um computador superou um campeão mundial em um jogo de xadrez (Russell; Norvig, 2013). Já, a IA forte ainda é objeto de grandes debates e pesquisas, visto que através dela o sistema seria capaz de realizar tarefas intelectuais sem intervenção humana.

Os sistemas de ASI são significativamente mais proficientes na execução de tarefas cognitivas humanas em comparação com as IAs, sendo trilhões de vezes mais inteligentes que os seres humanos em todos os aspectos (Webb, 2020). A ideia de uma inteligência tão avançada sugere que tais sistemas não apenas superariam os seres humanos em tarefas isoladas (como acontece com a IA atual), mas atingiriam um nível de autonomia e raciocínio capaz de transformar radicalmente a própria noção de inteligência.

Nessa perspectiva, “os sistemas projetados e construídos com técnicas de IA podem ser considerados como Sistemas de Comportamento Inteligente [...]” (Zamberlan et al., 2021). A inteligência exibida por esses sistemas está diretamente ligada às técnicas e algoritmos que os sustentam, sendo sempre um reflexo das regras e padrões estabelecidos por seus desenvolvedores. A inteligência a princípio, até então, não é autônoma no sentido absoluto, embora possam exibir comportamentos complexos e sofisticados.

Tais sistemas devem conter: a) uma base de conhecimento, um repositório abrangente de experiências, situações ou casos, organizados como fatos e regras ao longo de uma linha do tempo; b) um motor de raciocínio automatizado, apto a realizar deduções ou induções lógicas sobre as experiências arquivadas; c) um mecanismo que, por meio de treinamento reiterado, permita ao sistema identificar padrões intrínsecos

na base de conhecimento, promovendo, dessa forma, o aprendizado de máquina (Zamberlan et al., 2021).

Figura 01: Mapa mental ilustrando o funcionamento do sistema de comportamento inteligente.



Fonte: Autora e Coorientador (2024).

Ao observar a figura, percebe-se que o conceito de Inteligência Artificial (IA) é desmembrado em técnicas, algoritmos, ferramentas e metodologias. E, dentro do sistema de comportamento inteligente, destacam-se várias capacidades como o processamento de língua natural (PLN) e a visão computacional. O Processamento de Linguagem Natural (PLN), ou *Natural Language Processing* (NLP) em inglês, destaca-se como um dos campos tecnológicos de maior promessa e potencial. Essa subárea da inteligência artificial confere aos *softwares* a habilidade de ler documentos de forma inteligente, permitindo que os computadores compreendam, interpretem e manipulem a linguagem humana (Lage, 2020).

Por outro lado, a visão computacional é uma tecnologia que visa identificar objetos por meio de uma câmera, imitando o processo de percepção visual humana (Lage, 2020). São duas capacidades que se complementam no funcionamento dos *chatbots*, como no caso do ChatGPT, que utiliza essas tecnologias para a leitura de textos em formatos como PDF e a interpretação de imagens em JPEG. Esses sistemas também envolvem armazenamento de conhecimento para a manutenção de dados e informações, raciocínio automatizado, que pode ser dedutivo (inferindo

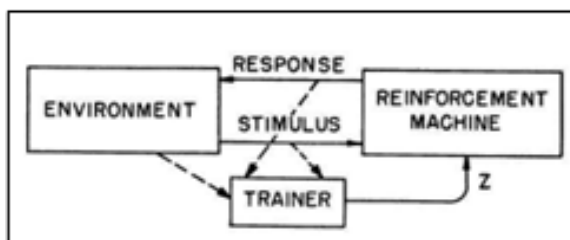
a partir de regras e premissas) ou indutivo (inferindo a partir de padrões observados), e o aprendizado de máquina, que reconhece padrões e melhora continuamente com base em dados.

No *machine learning*, o computador é projetado para “se autoprogramar” com base em sua própria experiência, coletando dados, interpretando informações e tomando decisões de maneira semelhante aos seres humanos (Nunes; Andrade, 2023). Esse tipo de aprendizado possui 03 subdivisões principais que são: supervisionado, não supervisionado e por reforço. Na aprendizagem supervisionada, o agente observa exemplos de pares de entrada e saída, e aprende uma função que realiza o mapeamento da entrada para a saída (Russell; Norvig, 2013). No aprendizado supervisionado, a máquina é treinada utilizando dados devidamente “rotulados”/“marcados” (Lage, 2020). Para clarificar, para ensinar um sistema a diferenciar maçãs de laranjas, é utilizado um conjunto de dados (*datasets*) composto por imagens rotuladas com atributos como cor, formato e textura. Dessa maneira, o modelo é treinado para associar essas características aos rótulos fornecidos.

Na aprendizagem não supervisionada, o agente identifica padrões na entrada sem receber qualquer *feedback* explícito e a tarefa mais comum nessa abordagem é o agrupamento, que consiste na identificação de grupos de exemplos de entrada que possam ser úteis (Russell; Norvig, 2013). Como não há um programador humano para fornecer instruções, o sistema aprende de forma autônoma a partir dos dados disponíveis (Lage, 2020, p. 70). Essa abordagem é amplamente utilizada para desenvolver sistemas de recomendação no Spotify e na Netflix, visto que com base em características e comportamentos verificados é possível personalizar sugestões. A título de exemplo, um usuário que assistiu à trilogia “*Matrix*” pode receber como recomendações os filmes como “O Exterminador do Futuro” (Cameron, 1984) e “*Blade Runner*” (Scott, 1982), pois os sistemas identificam o gênero de ficção científica e as semelhanças nos temas, como distopias futurísticas e questões relacionadas à inteligência artificial e ao destino da humanidade.

Pode-se afirmar que a aprendizagem por reforço abrange toda a IA: um agente é inserido em um ambiente e deve aprender a agir com sucesso nesse contexto e na aprendizagem por reforço, o agente aprende com base em uma sequência de reforços, recompensas ou punições” (Russell; Norvig, 2013). Logo, vê-se que o aprendizado se dá através do erro ou do acerto, pois não há interferência humana e nem base de dados como auxílio.

Figura 02: Esquema conceitual relacionado ao processo de aprendizagem por reforço.



Fonte: Minsky (1961).

Em diversos domínios complexos, a aprendizagem por reforço é o único método viável para treinar um programa com alto desempenho (Russell; Norvig, 2013). Esse modelo é frequentemente utilizado em jogos como Xadrez, Gamão e Go, pois esses ambientes estruturados permitem simulações extensivas, possibilitando que agentes explorem diferentes estratégias. Mais recentemente, pode ser encontrado, por exemplo, em veículos autônomos, que, por meio de um sistema sofisticado de sensores, são capazes de navegar por estradas urbanas e rurais, identificando obstáculos, seguindo placas e realizando diversas outras tarefas (Lage, 2020).

Destaca-se que, para implementar essa “heurística básica de aprendizagem”, é necessário generalizar as experiências anteriores, e uma maneira de se conseguir isso é por meio de modelos de decisão que se fortalecem com base no sucesso e ao empregar dispositivos que aprendem a identificar quais eventos estão associados ao reforço (recompensa), torna-se possível construir sistemas de “reforço secundário” com maior autonomia (Minsky, 1961). Ao aprender, não basta memorizar situações específicas, pois o sistema precisa identificar padrões e aplicar o que aprendeu a situações novas, mesmo que não sejam exatamente iguais às que já foram vivenciadas.

Todavia, ao aplicar esses métodos em problemas complexos, torna-se difícil distribuir de forma adequada o crédito pelo êxito de uma estratégia sofisticada entre as diversas decisões que a compõem (Minsky, 1961). Em uma partida complexa, o sucesso não vem de uma única jogada, mas sim de uma sequência de decisões interligadas, onde cada movimento influencia o resultado dos próximos. Essa situação exemplifica o que acontece quando aplicamos métodos de aprendizado por reforço em problemas complexos, dado que quando o sucesso é alcançado é difícil determinar exatamente qual ação ou grupo de ações foi responsável por esse resultado positivo.

O aprendizado profundo (*deep learning*) representa uma abordagem moderna e sofisticada dentro dos algoritmos de aprendizado de máquina, empregando múltiplas camadas para extrair, de maneira gradual, características de nível superior a partir dos dados brutos (Nunes; Andrade, 2023). Essa técnica consiste, essencialmente, em uma estratégia algorítmica que se fundamenta no conceito de redes neurais que a inspiração advém da compreensão da biologia dos cérebros humanos - as interconexões entre os neurônios (Lage, 2020). Nas primeiras camadas, a rede pode identificar elementos básicos, como bordas, cores ou formas simples em imagens, ou padrões sonoros em sinais de áudio. À medida que os dados passam por camadas sucessivas, essas informações básicas são combinadas para formar conceitos mais complexos, como o reconhecimento de objetos, rostos ou a compreensão de uma linguagem.

Ao contrário do cérebro humano, no qual um neurônio pode se conectar a qualquer outro, mesmo a uma certa distância física, as redes neurais artificiais são estruturadas de forma mais organizada, com camadas distintas, conexões predefinidas e direções específicas para a propagação dos dados (Lage, 2020). No cérebro humano, os neurônios estabelecem conexões altamente flexíveis e adaptáveis, formando uma rede interligada onde qualquer neurônio pode se conectar a outro, independentemente da distância física.

Essa funcionalidade é viabilizada pela presença de dois mecanismos de sinalização: a condução axônica e a transmissão sináptica (Oliveira, 1994). A condução axônica se refere ao movimento dos impulsos elétricos ao longo do axônio de um neurônio, permitindo que a informação percorra grandes distâncias dentro do sistema nervoso. Já, a transmissão sináptica é o mecanismo pelo qual a informação produzida ou processada por um neurônio é repassada a outro neurônio ou a uma célula efetora (Oliveira, 1994).

Imaginemos, por um instante, que a mente humana não é um simples repositório de informações, mas uma rede complexa de conexões entre percepções, emoções e experiências. Em muitas formas de aprendizagem, o que fazemos é perceber padrões e, a partir deles, ajustar nossa compreensão do mundo. As redes neurais artificiais buscam, com o auxílio de camadas sucessivas de processamento, realizar exatamente esse tipo de tarefa. Nessa perspectiva, as redes neurais artificiais de aprendizado profundo, também chamadas simplesmente de redes neurais, constituem um sistema projetado para operar de maneira semelhante ao cérebro e essas redes são organizadas

em camadas de neurônios matemáticos, permitindo o processamento de dados, a compreensão da fala humana e o reconhecimento visual de objetos (Lage, 2020). A informação é transmitida por meio dessas camadas, de modo que a saída de uma etapa serve como entrada para a próxima (Lage, 2020).

A primeira camada de uma rede é denominada camada de entrada, enquanto a última é designada como camada de saída e as camadas que ficam entre essas duas são chamadas de camadas ocultas (Lage, 2020). Por exemplo, uma imagem pode ser representada por uma matriz de pixels, onde cada pixel tem um valor numérico que descreve a intensidade da cor. Contudo, essa representação inicial não tem muito significado por si só; é apenas uma coleção de números. A camada de entrada, então, transmite esses números para a primeira camada oculta, onde começa o processamento real e refina as informações, tornando as representações cada vez mais detalhadas. O objetivo dessas camadas intermediárias é criar uma rede de conexões que consiga “aprender” com os dados.

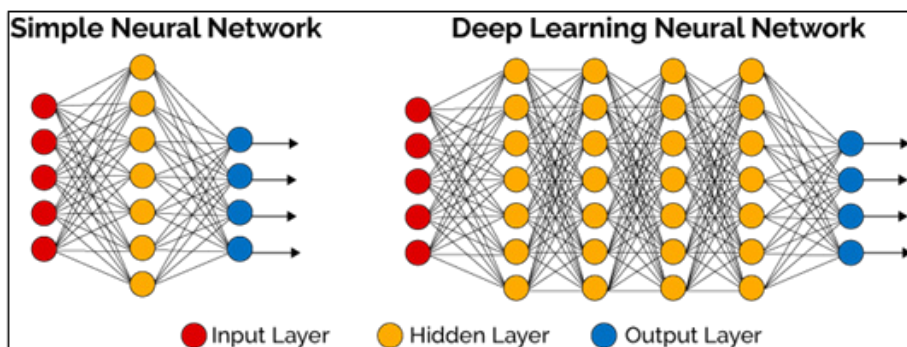
Finalmente, após esse processo de transformação e aprendizado, os dados chegam à camada de saída, onde a rede toma uma decisão ou faz uma previsão com base nas informações processadas. Se estivermos lidando com uma rede neural para reconhecimento de imagens, por exemplo, a camada de saída pode fornecer a classificação final, como “gato” ou “cachorro”.

A propósito, devido ao caráter ‘tudo ou nada’ da atividade nervosa, os eventos neurais e suas interações podem ser analisados por meio da lógica proposicional” (McCulloch; Pitts, 1990). A caracterização binária é o que torna a atividade nervosa comparável à lógica proposicional, onde as proposições também assumem valores binários de verdadeiro ou falso. Verifica-se que o comportamento de cada rede pode ser descrito nos termos mencionados, com a adição de componentes lógicos mais complexos para redes que incluem circuitos; e que, para qualquer expressão lógica que atenda a determinadas condições, é possível encontrar uma rede cujo comportamento corresponda à descrição proposta (McCulloch; Pitts, 1990).

Em síntese, enquanto o aprendizado de máquina emprega algoritmos para examinar dados, assimilar conhecimento desses dados e tomar decisões fundamentadas com base nesse aprendizado, o aprendizado profundo (*deep learning*) organiza algoritmos em camadas para formar uma “rede neural artificial” capaz de aprender e tomar decisões autônomas de maneira inteligente (Nunes; Andrade, 2023). A propósito, redes neurais

profundas podem ser mais difíceis de interpretar, tornando o processo de tomada de decisão do modelo uma “caixa pandora”, o que é uma preocupação quando se trata de transparência e confiança.

Figura 03: Distinção entre a arquitetura da rede neural simples e da rede neural no aprendizado profundo (*Deep learning*).



Fonte: Vázquez (2017).

Conforme a imagem ilustra, a Camada de Entrada (*Input Layer*), representada pelos nós vermelhos, recebe os dados iniciais, como características de informações (imagens, sons, texto, etc.). No centro, encontram-se as Camadas Ocultas (*Hidden Layers*), simbolizadas pelos nós amarelos, que processam os dados, aplicando funções matemáticas para identificar padrões ou relações entre as entradas. À direita, a Camada de Saída (*Output Layer*), com nós azuis, apresenta os resultados, como classificações.

Enquanto o modelo simples tem uma única camada oculta, redes neurais profundas incluem quatro camadas ocultas. Os nós são tipicamente organizados em uma série de camadas que são aproximadamente análogas a diferentes centros de processamento no córtex. Os computadores de hoje podem lidar com redes de “aprendizado profundo” com dezenas de camadas. Os nós são geralmente dispostos em uma sequência de camadas que podem ser comparadas aos diferentes centros de processamento no córtex (Lage, 2020).

Os computadores atuais são capazes de processar redes de “aprendizado profundo” compostas por dezenas de camadas (Lage, 2020). Diga-se de passagem que a empresa onde Neo trabalhava se chamava “Metacortex”, cujo nome, etimologicamente, deriva do grego *meta* (além ou após) e *córtex* (casca ou camada externa do cérebro). A nomenclatura

parece evocar a ideia de transcendência dos limites convencionais, sugerindo uma busca por uma realidade mais profunda e significativa.

No estado da arte/atual, a IA abrange uma ampla gama de subcampos, desde áreas gerais, como aprendizagem e percepção, até tarefas específicas, incluindo jogos de xadrez, demonstração de teoremas matemáticos, composição poética, condução de veículos em vias movimentadas e diagnóstico de enfermidades (Russell; Norvig, 2013). Essa diversidade de subcampos ilustra o avanço significativo da IA, que já desempenha um papel relevante em muitas esferas da sociedade, mostrando sua versatilidade e potencial.

A inteligência artificial de nível humano será alcançada, mas é quase certo que novas ideias serão necessárias, portanto, não é possível prever uma data com certeza, pode ser em cinco anos ou em quinhentos anos (inclinam-se apostas para que isso ocorra ainda no século XXI) (McCarthy, 2007). Neste ano de 2025, promete-se que a inteligência artificial evolua de maneira exponencial, alcançando avanços significativos em um único ano, possivelmente transformando radicalmente as capacidades tecnológicas e sociais, porém esse avanço técnico simboliza um abismo ontológico porque há o questionamento sobre o limite do que se entende sobre o que é “inteligente” e “humano”, por conseguinte, sobre o sentido da própria existência. Não é surpreendente que desenvolver uma IA comparável à inteligência humana seja tão difícil e que os avanços aconteçam lentamente, mesmo diante de conquistas importantes (McCarthy, 2007).

Fei-Fei Li relata que ao observar o progresso acelerado da inteligência artificial nos últimos 20 anos, especialmente nos subcampos de Aprendizado de Máquina (ML), Processamento de Linguagem Natural (NLP) e Visão Computacional (CV), percebe-se que os dados baseados na Web desempenham um papel crucial como força propulsora para o fortalecimento contínuo da IA (Yi Zeng; Ling Wang, 2017). Além disso, a cientista reconhecida como madrinha da IA afirma que é possível que os avanços na neurociência estimulem a interpretação da IA, e a ciência cognitiva represente outra força propulsora para acelerar esse processo (Yi Zeng; Ling Wang, 2017). O estudo do funcionamento do cérebro humano fornece subsídios para o desenvolvimento de modelos computacionais mais sofisticados, que buscam não apenas processar informações, mas também interpretar contextos e aprimorar a tomada de decisões.

Na verdade, a colaboração estreita entre pesquisadores de inteligência artificial e neurocientistas tem se mostrado um fenômeno emergente

na China, sendo o Centro CAS de Excelência em Ciência do Cérebro e Tecnologia de Inteligência (CEBSIT) o exemplo mais emblemático e nesse Centro adota uma estratégia que consiste em selecionar e recrutar os principais especialistas para enfrentar os desafios da interdisciplinaridade entre a ciência do cérebro e a inteligência artificial inspirada no cérebro (Zeng; Wang, 2017). Ao congregar cientistas de renome nas áreas de neurociência e IA, longe de ser uma simples soma de competências, configura-se como um esforço conjunto que visa não apenas o avanço da IA, mas a reinterpretação dos próprios limites da inteligência humana, mediada pela tecnologia. Por esse motivo Li expressa sua expectativa de que, em nível global, se intensifique a cooperação interdisciplinar em todas as áreas (Zeng; Wang, 2017).

A engenharia, medicina, negócios, arquitetura e pintura não estão preocupadas com o necessário, mas com o contingente; não com o que as coisas são, mas com o que elas poderiam ser, ou seja, com o design e a possibilidade de criar uma ciência ou ciências do design é tão grande quanto a possibilidade de criar qualquer ciência do artificial (Simon, 1996). A cooperação interdisciplinar proposta por Li se alinha diretamente com esse pensamento, sugerindo que, para avançarmos de maneira significativa, é fundamental que as áreas do conhecimento interajam de forma mais fluida e integrada. Desse modo, o design deixa de ser uma mera prática de criação estética, tornando-se uma ferramenta essencial para a transformação do mundo, seja no plano acadêmico, tecnológico, científico ou artístico.

O *design*, sob essa ótica, constitui o núcleo essencial de toda formação profissional, sendo a característica distintiva que separa as profissões das ciências, visto que escolas de engenharia, assim como aquelas de arquitetura, negócios, educação, direito e medicina, estão profundamente voltadas para o processo de design (Simon, 1996). Isso denota que em cada uma dessas áreas, o *design* é uma habilidade essencial, que não se limita a fazer projetos ou criar coisas, mas envolve a capacidade de pensar estrategicamente e de encontrar soluções criativas e eficazes para os desafios de cada profissão.

Superar a “fragilidade” dos sistemas de IA atuais e alcançar uma inteligência artificial de nível humano exige o desenvolvimento de programas capazes de lidar com problemas típicos de situações informáticas, onde as características relevantes para atingir um objetivo não estão previamente determinadas (McCarthy, 2007). A criação de uma IA verdadeiramente humana exigirá e apenas o desenvolvimento de algoritmos mais sofisticados, mas uma nova abordagem que permita a compreensão e com o mundo

à sua volta, sem depender de parâmetros previamente estabelecidos. Por ocasião, já em 1970, alguns defendiam a regulamentação internacional da IA, porém tal proposta foi uma má ideia na época e continuará sendo enquanto ainda tivermos tão pouco entendimento sobre como a IA de nível humano se manifestará (McCarthy, 2007, p. 1181).

Nesse diapasão, revela-se um cenário pós-apocalíptico, como o descrito no filme, em que máquinas dominam a terra utilizando os seres humanos como fonte de energia mantendo-os em cápsulas adormecidas – os chamados “campos de humanos cultivados”, não se distancia tanto da realidade contemporânea, em que redes sociais transformam a sociedade em um coletivo adormecido, facilmente manipulável por algoritmos projetados para influenciar decisões e comportamentos. Essa dependência tecnológica representa um tipo de “recarga” a favor das máquinas, na qual as capacidades humanas são exploradas como fonte de dados, atenção e engajamento. Diga-se de passagem que a robótica se dedica ao estudo de agentes inteligentes capazes de interagir e manipular o mundo físico (Russell; Norvig, 2013).

No entanto, há a possibilidade de inverter essa lógica: que as máquinas sejam recarregadas a serviço da humanidade. Isso requer regulamentações éticas que priorizem o fortalecimento da autonomia humana, utilizando a IA como ferramenta para resolver desafios globais. Assim, a “recarga” pode ser tanto um instrumento de emancipação quanto de dominação. O caminho escolhido depende das decisões tomadas agora, que moldarão o futuro da convivência entre seres humanos e máquinas.

2.2 Reestruturando a matrix: principais regulamentações da IA como recarga do sistema

Como já denunciava Pitty na canção “*Admirável Chip Novo*” (2003), vivemos em um mundo onde nada é genuinamente orgânico, visto que tudo parece previamente programado, estruturado para funcionar dentro de padrões impostos. E, diante de qualquer sinal de ruptura ou desajuste, aqueles que detêm o controle retornam para reinstalar o sistema em seu favor.

Em meio a um sistema pulsante, onde códigos invisíveis moldam cada interação e decisão, as regulamentações de Inteligência Artificial aparecem como uma promessa de “recarga do sistema”, uma tentativa de corrigir falhas, trazer equilíbrio e assegurar que a evolução tecnológica

sirva à humanidade. Todavia, se mal projetada, ela pode criar uma ilusão de domínio, um verniz de justiça que apenas encobre um sistema repleto de imperfeições, perpetuando desigualdades e reforçando estruturas que deveriam ser desmanteladas.

Dessa forma, este capítulo adentra esse universo de promessas e armadilhas, explorando como as regulamentações globais de IA procuram reprogramar as engrenagens do sistema. Tendo em vista que entre promessas e incertezas, o movimento de reestruturação avança, navegando entre a possibilidade de um futuro mais justo e o risco de criar um simulacro que apenas reforça o que já existe. Trazendo a baila a seguinte indagação: será o sistema realmente reprogramado em benefício para humanidade ou apenas recoberto por uma nova camada de controle?

Por isso, é crucial debater a transformação tecnológica à luz do processualismo constitucional democrático, a fim de garantir o estudo e a implementação das novas tecnologias e do *design* comportamental, visando criar arquiteturas de escolha e desenvolver respostas normativamente corretas, capazes de potencializar o acesso à justiça e mitigar as desigualdades sociais e jurídicas que persistem em nossa sociedade (Nunes, 2023a). Em outros termos, busca-se que a transformação digital reforce a justiça e a equidade, proporcionando um sistema que não apenas se adapte aos tempos modernos, mas que também avance na promoção de um acesso mais justo e universal à justiça.

Nesse ambiente, as abordagens regulatórias criadas podem ser classificadas em diferentes categorias, sobressaindo-se três principais: a regulação baseada em princípios, a regulação baseada em riscos e a regulação baseada em direitos (Filho, 2025). Cada abordagem apresenta vantagens e desafios e sua efetividade dependerá não apenas do modelo adotado, mas da capacidade do ordenamento jurídico de interpretá-las e aplicá-las de forma coerente e adaptável às transformações tecnológicas e sociais.

A regulamentação orientada pelos direitos humanos e fundamentais focaliza a proteção desses valores no contexto dos sistemas de inteligência artificial (Filho, 2025). Essa abordagem define, de forma antecipada, um conjunto específico de direitos e deveres a serem observados no desenvolvimento e na aplicação da IA, assegurando, sobretudo, a preservação dos direitos dos usuários (Filho, 2025). Isso implica que, antes da implementação das tecnologias, devem ser estabelecidas normas específicas que assegurem o respeito aos direitos dos indivíduos. Tal abordagem, longe de ser reativa, configura-se como preventiva.

A regulação baseada em princípios orienta o desenvolvimento e a aplicação de sistemas de inteligência artificial (IA) por meio da definição de premissas gerais e diretrizes éticas, dispensando a necessidade de estipular, de antemão, um conjunto rígido de direitos e obrigações (Filho, 2025). É uma abordagem que estabelece um arcabouço normativo flexível sem a rigidez de regras pré-determinadas. Diferentemente da baseada em princípios, marcada pela flexibilidade e pela orientação por valores gerais, as diretrizes normativas baseadas em riscos adotam uma abordagem mais estruturada e direcionada (Filho, 2025). Esse modelo categoriza os sistemas de inteligência artificial conforme seus níveis de risco, estabelecendo padrões de governança proporcionais à gravidade dos riscos identificados (Filho, 2025).

Exemplificada pelo *AI Act* da União Europeia e pelo Projeto de Lei nº 2338/2023 no Brasil, esse tipo de regulamentação busca equilibrar inovação e segurança, estabelecendo diretrizes proporcionais à periculosidade dos sistemas de inteligência artificial. O EU *AI Act* adota uma abordagem fundamentada na avaliação de riscos, classificando os sistemas de inteligência artificial em distintas categorias, tais como risco inaceitável, elevado, limitado e mínimo (Filho, 2025). Não há apenas o estabelecimento de requisitos legais rigorosos para sistemas de alto risco tendo em vista que aborda aspectos como governança de dados, transparência e segurança, mas também incorpora um ambiente regulatório experimental, conhecido como *sandbox* regulatório² (Filho, 2025).

Em 2021, a União Europeia (UE) apresentou uma proposta de regulamento do Parlamento Europeu e do Conselho que define regras harmonizadas sobre inteligência artificial, pois é do interesse da União preservar a liderança tecnológica assegurar que os cidadãos europeus possam desfrutar dos avanços tecnológicos em consonância com os valores, direitos fundamentais e princípios da União (UE, 2021). Ao instituir esse quadro jurídico uniforme, a União Europeia evidencia uma abordagem prudente, que não apenas antevê as potencialidades da inteligência artificial, mas também se debruça sobre os desafios que sua implementação acarreta para a sociedade.

2 O *sandbox* é uma solução flexível e estratégica que oferece um ambiente controlado onde inovações tecnológicas podem ser experimentadas sem a necessidade de uma implementação imediata e irreversível das regulamentações. Essa estratégia busca assegurar que a regulamentação seja adaptável e equilibrada, estimulando a inovação, enquanto protege os direitos dos usuários e a segurança (Filho, 2025).

No contexto do risco inaceitável, a União Europeia proíbe o uso de sistemas de inteligência artificial que constituam ameaça à segurança, aos direitos ou aos meios de subsistência das pessoas, e essa proibição abrange, especialmente, aplicações relacionadas à manipulação cognitivo-comportamental, ao policiamento preditivo, ao reconhecimento de emoções em ambientes de trabalho e instituições de ensino, assim como à classificação social (UE, 2025). A premissa de que a inovação não deve prosperar a qualquer custo fica clara, visto que ao estabelecer limites éticos e legais convoca os desenvolvedores e usuários a repensar seus paradigmas de atuação. Busca-se soluções que harmonizem o progresso tecnológico com o respeito à integridade humana.

Os sistemas de inteligência artificial classificados como de alto risco, tais como os utilizados em diagnósticos de enfermidades, na condução autônoma e na identificação biométrica de indivíduos envolvidos em atividades criminosas ou investigações penais, devem satisfazer um conjunto rigoroso de requisitos e obrigações para serem introduzidos no mercado da União Europeia (UE, 2025). Ao requerer processos de auditoria e verificação, a UE não apenas busca minimizar falhas operacionais, mas também prevenir possíveis vieses e discriminações inerentes aos algoritmos, o que é crucial em contextos sensíveis como o diagnóstico de doenças ou a identificação de indivíduos em investigações criminais.

Os sistemas de inteligência artificial classificados como de risco limitado, tais como robôs de conversação e geradores de conteúdo, devem observar obrigações de transparência, incluindo a exigência de informar os usuários sobre a origem automatizada dos conteúdos produzidos, permitindo que os indivíduos possam tomar decisões informadas quanto ao uso dessas tecnologias (UE, 2025). A clareza sobre a autoria dos conteúdos gerados por IA se torna essencial para evitar confusões entre produções humanas e automatizadas, pois em um contexto de disseminação de desinformação e de manipulação digital, saber que um texto, uma imagem ou um vídeo foi criado por um sistema artificial permite ao usuário adotar uma postura crítica e avaliar o conteúdo com o devido discernimento. Já, os sistemas de risco mínimo apresentam baixa periculosidade e por esse motivo não são regulamentados pelo regulamento da UE.

Em 24 de janeiro de 2024 o Jornal Oficial da União Europeia noticiou que a Comissão Europeia que criou o Serviço Europeu para a Inteligência Artificial, o qual, conforme o artigo 7º, alínea, C/2024/1459, deve colaborar para a promoção da cooperação internacional em relação

à regulamentação e governança da IA (UE, 2024). Nesse ínterim, em 21 de maio do corrente ano o Regulamento Europeu 2024/1689, também conhecido como *AI Act*, foi formalmente adotado pelo Conselho Europeu e entrou em vigor no dia 01 de agosto do mesmo período.

Desse modo, o *AI Act* leciona que “a IA é uma família de tecnologias em rápida evolução que contribui para um vasto conjunto de benefícios económicos, ambientais e sociais em todo o leque de indústrias e atividades sociais” (UE, 2024). Simultaneamente, devido às circunstâncias relacionadas à sua aplicação, utilização e grau de evolução tecnológica específicos, a inteligência artificial pode gerar riscos e comprometer interesses públicos, assim como direitos fundamentais protegidos pela legislação da União (UE, 2024). Possibilita observar, nas primeiras linhas da regulamentação, a linha tênue entre os benefícios e os riscos dessa tecnologia.

O *AI Act* inicia reconhecendo o imenso potencial da IA ao mesmo tempo em que alerta para os desafios decorrentes de sua aplicação. Diante da amplitude de seus efeitos sociais e da urgência em fortalecer a confiança pública, é imprescindível que a IA e o respectivo regime regulatório sejam desenvolvidos em consonância com os valores da União Europeia, conforme estabelecido no artigo 2º do Tratado da União Europeia (TUE), bem como com os direitos e liberdades fundamentais consagrados nos Tratados e, de acordo com o artigo 6º do TUE, com a Carta dos Direitos Fundamentais da União Europeia (UE, 2024).

Por isso, como premissa fundamental, a inteligência artificial deve ser concebida como uma tecnologia centrada no ser humano, servindo como um instrumento a seu serviço e tendo como propósito essencial a promoção do bem-estar da sociedade (UE, 2024). Ao delineá-la como uma ferramenta complementar, busca-se prevenir que se torne uma força descontrolada ou dissociada dos interesses sociais. É essencial adotar um regime jurídico da União que estabeleça normas harmonizadas relativas à inteligência artificial, com o objetivo de promover seu desenvolvimento, utilização e adoção no mercado interno, ao mesmo tempo em que assegure um elevado nível de proteção aos interesses públicos, como saúde, segurança e a defesa dos direitos fundamentais, incluindo a democracia, o Estado de direito e a proteção ambiental, conforme reconhecido e garantido pelo direito da União (UE, 2024).

Essas normas devem ser claras e robustas na proteção dos direitos fundamentais, incentivando novas soluções inovadoras e permitindo a formação de um ecossistema europeu composto por atores públicos

e privados que desenvolvam sistemas de inteligência artificial alinhados aos valores da União, ao mesmo tempo em que exploram o potencial da transformação digital em todas as regiões da União (UE, 2024). Para ser mais claro, a ideia é criar um modelo regulatório que, além de proteger os cidadãos, funcione como um motor de inovação, assegurando que o avanço tecnológico ocorra de forma ética e inclusiva.

O conceito de “sistema de IA” estabelecido no regulamento deve ser definido de maneira clara e precisa, estando alinhado de forma estreita com as iniciativas das organizações internacionais atuantes no campo da inteligência artificial. Isso visa garantir a segurança jurídica, promover a convergência internacional e assegurar uma ampla aceitação, ao mesmo tempo em que oferece a flexibilidade necessária para acompanhar as rápidas inovações tecnológicas nesse setor (UE, 2024). Então, há a necessidade de adaptabilidade semelhante à de um camaleão, que se ajusta ao seu ambiente de forma precisa.

A definição deverá ser fundamentada nas principais características dos sistemas de IA, que os diferenciam de sistemas de *software* ou métodos de programação tradicionais mais simples, e não incluirá sistemas baseados unicamente em regras estabelecidas por indivíduos para a execução automática de operações (UE, 2024). Trazendo à baila o que foi mencionado o tópico anterior, isto é, viabilidade de explorar a temática de maneira mais acessível, dispensando a necessidade de uma abordagem excessivamente complexa ou técnica. A ênfase recai sobre a possibilidade de se aprofundar no conteúdo com clareza e simplicidade, sem sacrificar a profundidade conceitual necessária para uma compreensão genuína do assunto.

Então, uma característica principal dos sistemas de IA é a sua aptidão para realizar inferências e essa capacidade diz respeito ao processo de gerar resultados, como previsões, conteúdos, recomendações ou decisões, que têm o potencial de influenciar tanto ambientes físicos quanto virtuais, e à capacidade dos sistemas de IA de gerar modelos, algoritmos ou ambos, a partir de entradas ou dados (UE, 2024). Como por exemplo, previsão de doenças, recomendações de conteúdos personalizados em plataformas de streaming, criação de algoritmos de otimização de rotas no trânsito com base em dados em tempo real.

As técnicas que possibilitam a realização de inferências na construção de um sistema de IA abrangem abordagens de aprendizagem automática, que utilizam dados para identificar maneiras de alcançar

objetivos específicos, e métodos baseados em lógica e conhecimento, que realizam inferências a partir de informações codificadas ou representações simbólicas da tarefa a ser solucionada (UE, 2024). A habilidade de um sistema de IA em fazer inferências ultrapassa o simples processamento de dados, viabilizando a aprendizagem, o raciocínio e a modelagem, desse modo, os objetivos do sistema de IA podem divergir da finalidade inicialmente prevista para ele em um contexto específico (UE, 2024).

Em outras palavras, a possibilidade de os resultados produzidos por um sistema de inteligência artificial se desviarem de seu propósito inicial pode ser ilustrada pelo uso de algoritmos de recrutamento automatizado. Quer dizer que um sistema projetado para selecionar candidatos com base em critérios de eficiência e competências técnicas pode, inadvertidamente, gerar resultados enviesados ao priorizar padrões ocultos nos dados históricos, como preferências de gênero ou etnia. Esse desvio compromete o objetivo original de promover um processo justo e imparcial. Com isso, para os fins deste regulamento, compreendem-se por “ambientes” os contextos nos quais os sistemas de inteligência artificial atuam, enquanto os resultados gerados por esses sistemas representam as distintas funções que desempenham, englobando previsões, conteúdos, recomendações ou decisões (UE, 2024).

Os sistemas de inteligência artificial são projetados para atuar com variados graus de autonomia, caracterizando-se por uma certa independência em relação às ações realizadas por intervenção humana e pela capacidade de operar de forma autônoma, sem a necessidade de interferência direta (UE, 2024). A independência propicia ganhos substanciais de eficiência, uma vez que os sistemas processam volumosos conjuntos de dados e executam atividades repetitivas com maior rapidez e precisão que os humanos. Não obstante, tal elevado grau de independência impõe desafios consideráveis no campo da segurança, pois sistemas altamente autônomos podem falhar.

Assim, para assegurar que os benefícios dos sistemas de inteligência artificial sejam plenamente alcançados, preservando os direitos fundamentais, a saúde, a segurança e o controle democrático, é indispensável promover a literacia em IA, capacitando fornecedores, responsáveis pela implementação e indivíduos afetados com os conhecimentos necessários para tomarem decisões informadas sobre esses sistemas e suas consequências (UE, 2024). A literacia em IA se refere ao entendimento e à capacidade de interpretar as tecnologias de IA de maneira crítica, compreendendo não apenas seu funcionamento, mas também as consequências éticas, sociais

e legais que podem advir do seu uso. De forma que a capacitação de todos os envolvidos facilita um ambiente mais seguro que promove o uso responsável da tecnologia.

Essas noções podem variar conforme o contexto em questão e englobam a compreensão da aplicação adequada dos aspectos técnicos durante o desenvolvimento do sistema de IA, as medidas a serem adotadas durante seu uso, as formas corretas de interpretar os resultados gerados pelo sistema de IA e, para as pessoas impactadas, os conhecimentos essenciais para entender como as decisões auxiliadas pela IA as influenciarão (UE, 2024). De modo que aqueles que desenvolvem esses sistemas devem possuir o conhecimento necessário para programá-los de maneira adequada e quando a IA é colocada em prática, é imperativo que seus operadores compreendam a forma correta de utilizá-la, assegurando sua eficácia e segurança. E os impactados pela aplicação devem entender como ela chegou ao resultado ou à decisão.

Embora a abordagem baseada no risco seja fundamental para a criação de um conjunto de regras eficazes e proporcionais, é crucial lembrar as Orientações Éticas para uma IA de Confiança, formuladas em 2019 pelo grupo independente IA HLEG, designado pela Comissão (UE, 2024). Nessas orientações, o IA HLEG elaborou sete princípios éticos não vinculativos para a IA, com o objetivo de assegurar que a inteligência artificial seja confiável e eticamente adequada, eles abarcam: iniciativa e supervisão por humanos; solidez técnica e segurança; privacidade e governação dos dados; transparência; diversidade, não discriminação e equidade; bem-estar social e ambiental e responsabilização (UE, 2024). Em termos simples, a avaliação de riscos funciona como um “mapa” que nos ajuda a identificar onde podem surgir problemas, como falhas técnicas ou questões de segurança, e a traçar estratégias para evitá-los. Por outro lado, as diretrizes éticas, representadas pelos sete princípios do IA HLEG, trazem uma perspectiva mais humana e social para o desenvolvimento da IA.

É igualmente essencial destacar que os titulares de dados preservam o pleno exercício de todos os direitos e garantias assegurados pelo direito da União, abrangendo, inclusive, aqueles relativos às decisões automatizadas individuais, como a definição de perfis (UE, 2024). O ato legislativo trata dos riscos inerentes à inteligência artificial, como vieses, discriminação e deficiências na responsabilização, ao mesmo tempo em que fomenta a inovação e incentiva a adoção dessa tecnologia (UE, 2024). Com essa

combinação é possível perceber uma preocupação em regulamentar e, ao mesmo tempo, incentivar a criação e a adoção de tecnologias de inteligência artificial, destacando a União Europeia como uma pioneira na tentativa de haver inovação e segurança.

Desse modo, como o primeiro ato legislativo mundial sobre supervisão da inteligência artificial, as normas da União Europeia têm o potencial de se tornar um marco de referência mundial tal como o Regulamento Geral de Proteção de Dados (RGPD) no campo da privacidade, promovendo uma inteligência artificial que seja ética, segura e confiável em escala global (UE, 2024). A influência do RGPD, que se consolida como padrão internacional para proteção de dados, é notável, especialmente ao servir de base para a criação da Lei Geral de Proteção de Dados (LGPD), Lei n. 13.709, que entrou em vigor em setembro de 2020 no Brasil.

Diga-se de passagem que a LGPD ganhou reconhecimento como a versão brasileira do RGPD, dada a profunda inspiração em suas diretrizes. A LGPD incorporou princípios essenciais, como transparência, segurança e responsabilização no tratamento de dados pessoais, sejam eles administrados por pessoas físicas ou jurídicas. Como finalidade a preservação e harmonização dos direitos fundamentais à liberdade e à privacidade, ao mesmo tempo em que estabelece limites para o consentimento de dados.

Nesse contexto, o Brasil, ao refletir o pioneirismo europeu, teria a oportunidade de mais uma vez se inspirar no novo capítulo da regulamentação europeia sobre inteligência artificial, delineando um marco normativo que equilibre inovação tecnológica, ética e proteção dos direitos fundamentais. Para mais, o estudo das propostas estrangeiras é crucial para a inserção do Brasil no contexto internacional, uma vez que os sistemas de IA estão propensos a ultrapassar fronteiras nacionais, por essa razão, é fundamental estabelecer normas brasileiras que estejam em conformidade com padrões técnicos e éticos alinhados aos parâmetros globais (Salomão, 2021).

O Senado brasileiro apresentou três projetos de lei: PL nº 5051/2019, o qual tem o intuito de definir os princípios para a utilização da IA no país (Brasil, 2019a); PL nº 5961/2019 que se refere à política em território nacional da IA nº 21/2020 (Brasil, 2019b); e o PL nº. 21/2020, o qual delineia princípios, direitos e obrigações para o uso da IA no Brasil e contempla outras disposições correlatas (Brasil, 2020). De maneira conjunta, esses projetos refletem uma postura que valoriza

não apenas o potencial inovador da inteligência artificial, mas também a importância de ancorá-la em uma base ética sólida. A justificação fornecida dentro deste Projeto para a regulamentação da matéria afirma claramente que os princípios, direitos e deveres estabelecidos serão obrigatórios, além de se propor a promover a expansão adequada da Inteligência Artificial no Brasil, com particular ênfase na segurança digital e na gestão de riscos (Salomão, 2021).

No dia 10 de dezembro de 2024 o Senado brasileiro aprovou projeto que regulamenta a IA no país e o texto aprovado é um substitutivo fundamentado no Projeto de Lei 2.338/2023 e o substitutivo também incorpora dispositivos propostos em outras sete iniciativas legislativas, incluindo o PL 21/2020, já aprovado pela Câmara dos Deputados. A versão aprovada nesta terça-feira excluiu os algoritmos das redes sociais da lista de sistemas classificados como de alto risco, mas o texto acolheu uma solicitação dos senadores governistas ao preservar o dispositivo que assegura a proteção dos direitos dos criadores de conteúdo e obras artísticas (Senado, 2024).

Esta Lei dispõe sobre diretrizes gerais de âmbito nacional para a governança responsável de sistemas de inteligência artificial (IA) no Brasil, visando à proteção dos direitos fundamentais, ao fomento da inovação responsável e da competitividade, bem como à garantia da implementação de sistemas seguros e confiáveis, promovendo o benefício da pessoa humana, a consolidação do regime democrático e o avanço social, científico, tecnológico e econômico (Brasil, 2024). Entretanto, a legislação não será aplicável aos sistemas de IA usados para fins pessoais sem objetivo econômico, nem aos desenvolvidos exclusivamente para defesa nacional. Também estão fora de seu escopo os sistemas em fase de pesquisa, testes ou desenvolvimento, desde que respeitem legislações específicas, como o Código de Defesa do Consumidor e a Lei Geral de Proteção de Dados.

O artigo 2º, inciso I, da legislação brasileira demonstra a mesma preocupação com a centralidade da pessoa humana presente no Regulamento Europeu sobre governança da inteligência artificial. Leva-se em conta o desenvolvimento, a implementação e o uso de sistemas de IA são sustentados por princípios que valorizam os direitos fundamentais, com destaque para a promoção dos direitos humanos e a defesa dos valores democráticos. Entre os pilares essenciais estão a liberdade de expressão, o livre desenvolvimento da personalidade, a proteção ao meio ambiente e o equilíbrio ecológico.

A igualdade, a não discriminação, a pluralidade e a diversidade são princípios indispensáveis, alinhados à valorização do trabalho humano e aos direitos sociais. Ademais, a inovação e o desenvolvimento socioeconômico, científico e tecnológico também figuram como prioridades, sendo complementados pela defesa do consumidor, pela livre iniciativa e pela livre concorrência. E a privacidade e a proteção de dados pessoais são temas centrais acompanhados do respeito à autodeterminação informativa. Com isso, elenca-se a promoção da pesquisa e do desenvolvimento reduzir desigualdades e estimular inovações tanto no setor público quanto no privado, incluindo parcerias público-privadas.

Ademais, sustenta-se que a integridade e a confiabilidade das informações são essenciais para garantir a liberdade de expressão e o acesso pleno à informação. Outros pilares incluem o fortalecimento do processo democrático e do pluralismo político, proteção de direitos autorais, propriedade intelectual e segredos comerciais, bem como a segurança da informação e a cibernética. Estando o Brasil comprometido à inserção e competitividade no mercado internacional, promovendo a cooperação global e se alinhando aos padrões técnicos e regimes de obrigações nacionais e internacionais.

O artigo 3º do projeto dispõe que o desenvolvimento, a implementação e o uso de sistemas de inteligência artificial devem se pautar na boa-fé, respeitando um conjunto de princípios. Entre os fundamentos que norteiam essas práticas, destaca-se o compromisso com o crescimento inclusivo, o desenvolvimento sustentável, autodeterminação e a liberdade de decisão e escolha. Menciona-se o combate a qualquer forma de discriminação ilícita ou abusiva, promovendo justiça, equidade e inclusão em todas as esferas.

Ainda, assegura-se que a supervisão humana seja efetiva e proporcional ao grau de risco envolvido em todo o ciclo de vida da IA, considerando que no projeto de lei brasileiro a categorização de riscos é dividida em duas: risco excessivo e alto risco. O artigo 13, que trata do risco excessivo, abrange sistemas de inteligência artificial cuja aplicação pode representar uma ameaça significativa à segurança e aos direitos fundamentais. Por sua vez, no artigo 14 a definição dos sistemas de alto risco é realizada com base na gravidade dos impactos que podem causar a indivíduos ou grupos afetados. Ele abrange diversas áreas, incluindo segurança, educação, trabalho, serviços essenciais, administração da justiça, saúde e controle de fronteiras.

A transparência e a explicabilidade dos sistemas de IA também figuram como princípios fundamentais, considerando o segredo comercial e industrial, bem como a participação de todos os agentes envolvidos na cadeia de valor. Somam-se a isso a diligência devida e a auditabilidade ao longo de todo o ciclo de vida do sistema, respeitando o estado da arte do desenvolvimento tecnológico e o grau de risco associado. A confiabilidade e a robustez dos sistemas de IA são essenciais para assegurar a proteção dos direitos e garantias fundamentais, incluindo o devido processo legal, a contestabilidade e o contraditório.

Nesse sentido, a prestação de contas, a responsabilização e a reparação integral de danos são pilares indispensáveis para uma governança ética e eficiente. Adicionalmente, a prevenção, a precaução e a mitigação de riscos e danos devem ser priorizadas, garantindo que os métodos empregados sejam proporcionais às finalidades legítimas do sistema. O desenvolvimento e o uso ético da IA, aliados a uma governança transparente e participativa, visam proteger direitos fundamentais individuais, sociais, coletivos e econômicos.

Nesse sentido, o artigo 4º delimita conceitos essenciais à compreensão e aplicação desta Lei, estabelecendo definições precisas que norteiam a interpretação de seus dispositivos e asseguram a uniformidade na sua implementação. O Sistema de inteligência artificial (IA) consiste no mecanismo automatizado que, com diferentes níveis de autonomia e com finalidades explícitas ou implícitas, processa dados ou informações recebidas para produzir resultados, tais como previsões, conteúdos, recomendações ou decisões, capazes de impactar ambientes virtuais, físicos ou reais (Brasil, 2024).

Já, o ciclo de vida reside no conjunto de etapas que abrangem desde a concepção, planejamento, desenvolvimento, treinamento, retreinamento, testagem, validação, implantação e monitoramento de um sistema de IA, incluindo possíveis modificações e adaptações, podendo sua descontinuidade ocorrer em qualquer uma dessas fases (Brasil, 2024). Reflete-se que não é um sistema operacional estático, mas um organismo em constante evolução, cujos efeitos se transformam a cada etapa.

O Sistema de inteligência artificial de propósito geral (SIAPG) é o sistema de IA baseado em modelos treinados com amplas bases de dados, concebido para desempenhar uma variedade de tarefas distintas e atender a múltiplas finalidades, mesmo aquelas para as quais não foi especificamente desenvolvido e treinado, sendo passível de integração em

diversos sistemas ou aplicações (Brasil, 2024). Gerando perquirições sobre sua previsibilidade e responsabilidade, visto que existindo a possibilidade desses sistemas serem utilizados por terceiros, de formas inesperadas, torna-se mais difícil antecipar suas ameaças.

Nessa toada, Inteligência artificial generativa (IA generativa) representa o modelo de IA projetado especificamente para criar ou modificar de maneira substancial, com diferentes níveis de autonomia, elementos como texto, imagens, áudio, vídeo ou código de *software* (Brasil, 2024). Diga-se de passagem que no texto original, a palavra “*software*” está em negrito. Cumprindo destacar a distinção entre hardware e software, uma vez que o primeiro, assim como o corpo humano, é a parte física e tangível que possui componentes como processadores, memória, e dispositivos de armazenamento.

Por seu lado, o segundo pode ser comparado à alma, pois representa o conjunto de programas, códigos e instruções que orientam o sistema. Então, o *hardware* é o corpo, a estrutura física, enquanto o *software* é a alma, que dá vida e direção ao funcionamento do sistema e ambos são indispensáveis para o funcionamento adequado de qualquer dispositivo ou sistema computacional. Importa ressaltar que se trata de uma analogia de caráter exclusivamente ilustrativo, cujo objetivo é apenas facilitar a compreensão, sem pretensão de corresponder a uma equivalência.

O desenvolvimento, a distribuição e a aplicação de sistemas de inteligência artificial (IA) envolvem diferentes papéis fundamentais, que são desempenhados por agentes distintos. Os agentes de inteligência artificial são desenvolvedores, distribuidores e aplicadores que atuam na cadeia de valor e na governança interna de sistemas de IA (Brasil, 2024). Reconhecendo que o desenvolvimento e a implementação da IA envolve múltiplas etapas e diferentes atores. Inclusive, seja dito de passagem que a Matrix é sustentada pela atuação de agentes responsáveis por solucionar falhas no sistema, como Morpheus.

O desenvolvedor, seja pessoa natural ou jurídica, é responsável pela criação do sistema de IA, seja de forma direta ou por encomenda, visando sua comercialização ou aplicação em serviços próprios, sob sua marca ou nome (Brasil, 2024). Ao reconhecer o ato de criação como um ato de autoria, que carrega em si um compromisso inescapável com os efeitos práticos e sociais dessa inovação. O desenvolvedor, ao criar a tecnologia e introduzi-la no mercado não apenas a materializa como um produto técnico, mas a reveste de valores e responsabilidades. Longe de ser um mero

formalismo, essa atribuição constitui um imperativo ético: toda inovação tecnológica deve se submeter a um crivo normativo que considere não apenas os avanços técnicos, mas também os impactos sociais, políticos e jurídicos decorrentes de sua aplicação.

Já o distribuidor atua como intermediário, disponibilizando o sistema de IA para que terceiros o utilizem, seja de forma onerosa ou gratuita, ampliando seu alcance (Brasil, 2024). É uma ponte entre o desenvolvedor do sistema e o usuário final. Para compreender melhor essa dinâmica, é possível fazer uma analogia com a comercialização de livros em que o autor escreve a obra (assim como o desenvolvedor cria o sistema de IA), mas para que o livro chegue às mãos dos leitores, ele precisa de um editor e de uma livraria que o distribuam. Da mesma forma, no caso da inteligência artificial, o distribuidor assume esse papel ao conectar a tecnologia aos usuários.

Por último, o aplicador é aquele que emprega o sistema em seu nome ou benefício, configurando-o, mantendo-o e fornecendo dados essenciais para sua operação e monitoramento (Brasil 2024). Diferente do desenvolvedor, que cria a tecnologia, e do distribuidor, que viabiliza seu acesso, o aplicador é quem efetivamente emprega a IA para alcançar determinados objetivos, seja em benefício próprio ou de terceiros. Esses papéis interdependentes garantem o ciclo de vida completo dos sistemas de IA, desde sua criação até sua utilização efetiva, sendo cruciais para a disseminação e implementação bem-sucedida dessa tecnologia.

A autoridade competente é a entidade da administração pública federal, dotada de autonomia técnica e decisória, que coordenará o Sistema Nacional de Regulação e Governança de Inteligência Artificial (SIA) (Brasil, 2024). A autonomia técnica e decisória atribuída a essa autoridade é fundamental, pois ela precisa ter a capacidade de tomar decisões informadas sobre a regulação da IA, sem pressões externas que possam comprometer a imparcialidade ou a eficácia de suas ações. Por seu turno, o SIA, visa fomentar a cooperação e a harmonização entre diversas autoridades setoriais, como as vinculadas à saúde, educação, finanças, entre outras, além de integrar outros sistemas nacionais que possam se envolver com a aplicação da inteligência artificial, como os referentes à proteção de dados pessoais e à defesa do consumidor.

O ambiente regulatório experimental (*sandbox regulatório*) é delimitado como um espaço temporário para testar inovações em sistemas de IA, dentro de critérios pré-estabelecidos, como no *AI Act* (Brasil, 2024).

Como o modelo proposto pelo *AI Act* na União Europeia é oportunidade de novas soluções tecnológicas serem experimentadas de maneira supervisionada, com regras claras que garantem a conformidade com princípios éticos e legais. Por sinal, no filme *Ex Machina* um programador é convidado a avaliar uma inteligência artificial avançada em um ambiente isolado, busca avaliar a capacidade da IA em demonstrar comportamento humano convincente (Teste de Turing) e ao passo que a trama se desenrola, emergem complexas questões éticas, como o livre arbítrio, o controle e o papel da IA em nossa sociedade.

E o estado da arte do desenvolvimento tecnológico é associado às práticas e meios técnicos mais avançados e baseados em evidências (Brasil, 2024). O termo estado da arte se refere ao estágio mais recente e refinado de uma determinada área, representando o ápice das inovações e descobertas já alcançadas. No contexto da IAI, isso se traduz nas soluções tecnológicas mais sofisticadas, como algoritmos de aprendizado profundo, redes neurais avançadas e outras metodologias que estão na vanguarda do desenvolvimento técnico.

Ao visualizar uma matriz comparada entre a legislação europeia e o projeto de lei brasileiro, vê-se que no que tange à quantidade de obrigações descritas, o PL 2.338/2023 mostra-se mais extenso que o marco europeu tendo em vista que são 56 obrigações contempladas na proposta brasileira, ante 39 previstas na Lei de Inteligência Artificial Europeia (ITS, 2024). O número de obrigações não se traduz, por si só, em um parâmetro absoluto de qualidade ou eficácia, porém a proposta brasileira, ao prever um conjunto mais extenso de obrigações, pode estar dialogando com um desejo de antecipar cenários de risco e estabelecer uma regulação mais minuciosa que abarque a multiplicidade de desafios e complexidades inerentes ao desenvolvimento e uso da IA no país.

Na atribuição de responsabilidades ao longo da cadeia de valor, o PL 2.338/2023 adota uma abordagem mais horizontal, distribuindo obrigações a todos os envolvidos, ao passo que o marco europeu segue uma estratégia mais direcionada, impondo deveres a atores específicos (ITS, 2024). Podendo ser interpretada como um reflexo da busca por uma responsabilidade compartilhada, em que cada ator, independentemente de sua posição na cadeia produtiva, carrega consigo um dever de diligência e de compromisso com os princípios fundamentais que orientam a regulação da IA, como a proteção de dados, a transparência e a segurança.

No marco europeu, poucas obrigações possuem aplicação universal, sendo mais comum sua incidência sobre o desenvolvedor (a quem se destinam 38 das 39 obrigações), seguido pelo distribuidor, que assume 12 dessas exigências (ITS, 2024). Entretanto, ao concentrar a responsabilidade em determinados agentes existe a possibilidade de gerar uma visão reducionista das complexas relações que envolvem o processo, visto que existem muitas interdependências presentes no ecossistema da IA.

No brasileiro, ao adotar a definição de “agentes da inteligência artificial”, estabelece a solidariedade entre fornecedor, aplicador e distribuidor em 33 das 56 obrigações previstas (ITS, 2024). Refletindo uma compreensão mais holística da responsabilidade que busca compreender um fenômeno em sua totalidade, considerando a interconexão entre suas diversas partes e os múltiplos fatores que influenciam seu funcionamento. Dessa forma, a regulação da inteligência artificial não deve ser encarada apenas como uma questão técnica ou jurídica, mas como um desafio filosófico, exigindo constante interpretação à luz dos direitos fundamentais.

Ademais, ressalta-se que o Brasil é um dos países signatários da “*Declaração de Bletchley*” que foi publicada no dia 01 de novembro de 2023, a qual identifica oportunidades, riscos fronteiriços associados à IA e ações para enfrentá-los, com a intenção de promover colaboração internacional e o desenvolvimento responsável nesse campo. A Declaração destacou que a Inteligência Artificial (IA) oferece vastas oportunidades em escala global, com o potencial para transformar e aprimorar o bem-estar humano, a paz e a prosperidade (Reino Unido, 2023).

Ao passo que nos Estados Unidos, assiste-se a uma transformação tecnológica sem precedentes, impulsionada pelas gigantes do setor como Google, Microsoft, Amazon e Meta. Dado que essas empresas têm liderado o desenvolvimento de inteligência artificial com uma velocidade impressionante, destacando-se globalmente. No entanto, essa aceleração tem sido acompanhada por uma regulação limitada e fragmentada, refletindo a tradicional resistência americana à intervenção governamental em mercados privados.

Embora o país tenha algumas iniciativas e regulamentos pontuais, como as orientações emitidas por órgãos como a *Federal Trade Commission* (FTC) e o *National Institute of Standards and Technology* (NIST), ainda não há uma legislação federal abrangente que estabeleça diretrizes claras e uniformes para a IA. Refletindo uma preocupação menor com os aspectos

éticos e de responsabilização, o que pode gerar desafios em termos de segurança, privacidade e justiça social. Nos Estados Unidos, embora o Congresso continue a discutir diversas propostas, os estados têm progredido de maneira independente, promulgando leis que tratam tanto de aspectos gerais quanto específicos no que tange à utilização de inteligência artificial (Diniz, 2024).

Em outubro de 2022, foi publicado o *Blueprint for an AI Bill of Rights*, um documento técnico concebido para orientar a elaboração de políticas voltadas à proteção dos direitos civis e à promoção dos valores democráticos, mas não possui caráter de legislação formal (Filho, 2025). O intuito é sensibilizar os desenvolvedores e os responsáveis pela implementação de tecnologias sobre a necessidade de garantir que essas ferramentas atendam aos princípios democráticos e aos direitos humanos, posto que o projeto foi elaborado através de consultas acadêmicas, grupos de defesa dos direitos humanos e entre outras organizações e empresas.

A estrutura do documento se baseia em cinco princípios fundamentais: segurança e eficácia dos sistemas; proteção contra discriminação algorítmica; privacidade de dados; aviso de notificação e explicação; e alternativas humanas e mecanismos de recurso (Krantz; Jonker, 2024). Esse movimento pode ser interpretado como um passo inicial para a construção de um arcabouço regulatório nos Estados Unidos, alinhado às tendências internacionais de governança da IA.

Em 30 de outubro de 2023, o presidente dos Estados Unidos, Joe Biden, promulgou uma ordem executiva com o intuito de estabelecer regulamentações para a inteligência artificial (IA) no território nacional. E nessa estrutura, os órgãos federais são responsáveis por emitir diretrizes baseadas em princípios e deixam a implementação para o setor privado e a *Federal Trade Commission* (FTC), por exemplo, publicou diretrizes que destacam a importância de sistemas de IA serem transparentes, explicáveis, justos e responsáveis (Filho, 2025). A promulgação dessa ordem executiva e as subsequentes diretrizes da FTC indicam um esforço crescente dos Estados Unidos para estabelecer uma governança robusta da IA.

O *Department of Defence* (DoD) adotou princípios éticos para IA aplicáveis a funções de combate e não-combate, focando em responsabilidade, equidade, rastreabilidade, confiabilidade e governabilidade (Filho, 2025). O esforço do DoD em estabelecer padrões éticos e mecanismos robustos de controle para o uso da IA não é apenas uma medida de segurança operacional, mas também uma estratégia para evitar que os avanços

tecnológicos se transformem em uma arma de destruição descontrolada. Ao antecipar possíveis cenários de batalha entre sistemas autônomos, busca-se garantir que a supervisão e a responsabilidade humana permaneçam no centro das operações.

Devido a crescente inquietação com a regulação da inteligência artificial, surgiram novas diretrizes legislativas específicas com o propósito de estabelecer parâmetros para o desenvolvimento e a utilização ética dessas tecnologias. Entre elas, o *Algorithmic Accountability Act* e o *Consumer Online Privacy Rights Act*, que visam criar regulamentações específicas para decisões automatizadas e privacidade online (Filho, 2025). Entretanto, no dia 20 de janeiro de 2025, o atual presidente dos Estados Unidos, Donald Trump revogou o texto promulgado pelo ex-presidente que estabelecia medidas para a proteção dos cidadãos norte-americanos contra os potenciais riscos dos sistemas de inteligência artificial, sob o argumento de que tais diretrizes impunham obstáculos ao desenvolvimento da inovação no setor. Dessa forma, o país segue sem uma legislação federal específica para a regulamentação da IA. Uma regulamentação excessiva pode inibir o avanço tecnológico e reduzir a competitividade internacional dos setores produtivos.

O fundador e CEO da Meta, Mark Zuckerberg, publicou em seu perfil no Instagram que a sua empresa se alinha ao projeto de direita promovido por Donald Trump e Elon Musk, mencionando-os explicitamente na publicação (Vilaça, 2025). Essa aproximação pode indicar uma busca por menos intervenção estatal e maior controle corporativo sobre o ecossistema digital, o que pode reforçar tanto o poder dessas grandes empresas quanto às tensões sobre sua influência política e social. Nesse sentido, as *big techs* são comparáveis às máquinas sencientes do filme *Matrix* (Vilaça, 2025). Dado que essas empresas exercem influência sobre a informação, os dados e até as percepções sociais, moldando o comportamento coletivo de maneira muitas vezes imperceptível.

Na realidade simulada a humanidade estava sob o domínio do Agente Smith, uma inteligência artificial projetada pelas máquinas para impedir que os indivíduos despertassem da ilusão imposta pelo sistema (Vilaça, 2025). Lança-se, desse forma, um alerta sobre os riscos inerentes às tecnologias, cuja força transformadora, embora inegável, pode também favorecer a alienação e a passividade. Nesse “novo mundo” um grupo de tecnocratas bilionários do Vale do Silício (que não são inteligências artificiais, mas os verdadeiros proprietários das máquinas) exerce um

controle semelhante ao dos agentes Smith porque atuam como guardiões da ordem digital, garantindo que a sociedade permaneça imersa em um estado de dormência induzido pelo excesso de estímulos, enquanto controlam os dois mundos de forma simultânea: a terra e a internet (Vilaça, 2025).

Em meio às regulamentações pioneiras, outras ainda em fase de aperfeiçoamento, e em contextos que mais se assemelham a uma terra sem lei, a inteligência artificial segue uma ascensão vertiginosa, ganhando cada vez mais espaço em todos os setores da sociedade, especialmente no judiciário. O ano de 2025 desponta com a promessa de especulação de que os modelos de linguagem conquistarão uma relevância ainda maior, moldando de maneira inusitada o futuro de diversas áreas. Assim, seguimos viagem com a Nabucodonosor para o próximo capítulo, rumo ao outro lado do mundo jurídico, onde a versão *hi-tech* da inteligência artificial ganha terreno.

3 O CENÁRIO ESTABELECIDO PELO PROCESSUALISMO (TECNO)NEOLIBERAL E OS DESAFIOS DA INTELIGÊNCIA ARTIFICIAL

“Você nasceu em uma prisão que não pode ver, sentir ou tocar. Uma prisão para sua mente.”

– Morpheus, em Matrix (Wachowski; Wachowski, 1999)

Preliminarmente, tem-se que para um entendimento profundo e sólido do direito processual civil contemporâneo, bem como da conduta judiciária nele estabelecida, é imperativo voltar a uma pré-compreensão da jornada do Estado Moderno, desde seus primórdios até o desenvolvimento do conceito político de Estado (Isaia, 2012). Dado que a perspectiva histórica surge como um esteio essencial para um discernimento completo e contextualizado dos desenvolvimentos normativos e institucionais que moldaram o cenário jurídico atual.

O advento do conceito político de Estado se materializou durante os séculos XVII e XVIII, uma época em que a convivência social parecia carecer de ordem, encontrando-se permeada por conflitos e desassossegos (Isaia, 2012). Diante de tais circunstâncias, observou-se o florescimento de manifestações distintas de projetos civilizatórios que, sobretudo, provocaram transformações nos alicerces políticos e sociais da sociedade europeia (Isaia, 2012). Com isso, imerso no estado de natureza, desponta-se o pacto social como uma engenhosa aspiração para transcender o estágio de conflito e alcançar a harmonia social.

A transposição do individualismo inerente ao estado de natureza levou à edificação de um novo panorama político encarnado pelo contrato social, cujo poder não reside unicamente na figura de um único indivíduo, mas sim em todo um organismo político fundamentado nos pilares da igualdade perante a vida, liberdade e propriedade (Isaia, 2012). Esse prisma promove uma redefinição das fundamentações da governança, visando o alcance de um equilíbrio entre os interesses diversos que permeiam a coletividade.

Diante disso, Locke, com a perspectiva em combater os abusos monárquicos, incitou a essencial divisão dos poderes (legislativo, executivo

e judiciário) buscando a constituição de representantes capazes de unir o corpo político sob um governo que assumisse a responsabilidade para com todos os membros da sociedade (Isaia, 2012). Desse modo, essa estrutura, delineada pelo filósofo francês, representa um dos pilares do constitucionalismo moderno, visto que caracteriza um governo democrático.

A essência primordial do Estado Liberal de Direito reside em sua característica de limitação jurídico-legal negativa, ou seja, sua função como garantia aos indivíduos-cidadãos diante de possíveis atuações estatais que restrinjam ou constringam sua liberdade de ação no cotidiano (Streck; Morais, 2014). Isto significa que, nessa visão, o Estado deve agir dentro dos limites estabelecidos pela lei e garantir os direitos individuais.

Entretanto, o paradigma do Estado Liberal se revelou insuficiente para atender às novas aspirações do povo. Assim sendo, sem desprezar as conquistas e os princípios fundamentais trazidos pelo liberalismo burguês, busca-se atribuir uma nova essência axiológico-política (Streck; Morais, 2014). Logo, o modelo liberal clássico cedeu lugar ao modelo social. O Estado Social de Direito delineia um modelo no qual o bem-estar e o progresso social assumem um papel central, norteando as ações do poder público (Streck, Morais, 2014).

Em outras palavras, em vez de descartar completamente as ideias importantes do liberalismo, começa-se a atribuir a elas um novo significado e aplicação. Essa mudança resultou em um governo que tem como aspecto central a garantia dos direitos sociais e a promoção do desenvolvimento coletivo da sociedade. Observa-se que há mudanças tanto na forma de agir do governo quanto em seus objetivos finais. O advento do Estado Democrático de Direito emerge em meio à crise do Estado Social, sendo uma manifestação desse histórico que reconhece a complexidade social. Então, o Estado Democrático de Direito vai além não apenas da ótica do Estado Liberal de Direito, mas também da do Estado Social de Direito (Streck; Morais, 2014).

O Estado Democrático de Direito transcende a esfera material de assegurar uma vida digna ao ser humano, adquirindo um papel simbólico como agente propulsor da participação pública no processo de construção e reconstrução de um projeto social, ele dissemina uma perspectiva futurista voltada para a criação de uma nova sociedade, na qual a questão da democracia abarca e implica, necessariamente, a resolução das questões fundamentais relacionadas às condições materiais de existência (Streck;

Morais, 2014). Em outros termos, a democracia desempenha uma função importante na solução de temas relacionados às condições de vida das pessoas, pois permite que todos tenham voz e participem das decisões que as afetam, garantindo que as necessidades e interesses de todos sejam considerados.

Nesse diapasão, ao contrário dos modelos preexistentes, o Estado Democrático de Direito, além de ser uma continuidade, manifesta-se como uma ruptura, uma vez que, a partir dos textos constitucionais direcionadores e compromissados, ele explicita tanto formal como materialmente as condições necessárias para a modificação da realidade (Streck; Moraes, 2014). É um sistema que estabelece diretrizes claras sobre o que precisa ser feito para melhorar determinado conjunto de circunstâncias. Até o advento do Estado Democrático de Direito (o chamado “Direito *pós-Auschwitz*”) o direito se restringia a uma racionalidade meramente instrumental (Streck, 2015).

Quando ele alcança elevado grau de autonomia, passa a poder funcionar como escudo contra agentes predadores, sobretudo aqueles exógenos, como a política, a moral e a economia (Streck, 2015). Então, teoricamente falando: a política deixa de subordinar as decisões judiciais a conveniências; a moral não impõe visões de mundo arbitrárias; e o capital econômico não consegue capturar o processo decisório em benefício exclusivo de grupos poderosos.

Isto posto, torna-se evidente que cada período brevemente mencionado teve suas peculiaridades e contribuições para o momento hodierno em que vivemos. Do mesmo modo que, é crucial para discutir a história do processo civil, pois ao considerar a evolução das ideias políticas e jurídicas é possível compreender como o processualismo também se desenvolveu ao longo dos séculos. Até aqui, buscou-se compreender o caminho percorrido até a consolidação do Estado Democrático em que vivemos. Agora, o próximo passo é compreender as raízes do processo civil.

Ao observarmos as intrincadas teias do direito processual civil contemporâneo, somos conduzidos a uma jornada ancestral que se inicia na Roma Antiga, mais precisamente no período justiniano (Isaia, 2012). Foi nesse período que nasceu a gênese do rito ordinário, marcado por uma distinção fundamental entre julgamento e decisão, como mencionado anteriormente. Em razão disso, o sistema jurídico passou a se estruturar em torno da codificação das leis, organizando-as em códigos e implementando uma abordagem processual civil “fase a fase”, esse

sistema foi consideravelmente ampliado com a ascensão do movimento racionalista-normativista, que exerceu uma forte influência durante o período de 1601 a 1700 (Isaia, 2012).

Em acréscimo, o Iluminismo defendeu a ideia de que a lei deveria ter um caráter determinado, com o objetivo de restringir a discricionariedade judicial e assegurar a segurança da liberdade civil, as leis devem ser claras e precisas, de modo a vincular inequivocamente a atuação do juiz. Ou seja, ao defender a aplicação rigorosa e objetiva das leis, o movimento iluminista buscou impedir que os juízes assumissem funções legislativas, garantindo a preservação da separação dos poderes e a imparcialidade do sistema jurídico.

No Brasil a influência racionalista foi refletida na elaboração do Código de Processo Civil, que seguiu uma abordagem lógica e formal, própria da mentalidade algébrica dos séculos XVIII e XIX (Isaia, 2012). Buscou-se uma organização meticulosa processual. Desse modo, o processo de conhecimento, em sua natureza declaratória e ordinária, corresponde a um procedimento no qual a verdade é proclamada ao final de um amplo debate judicial, resultando em um juízo de certeza formulado pelo magistrado com base na plena utilização dos meios de ataque e defesa disponibilizados às partes litigantes (Baptista da Silva, 2004).

No entanto, essa abordagem consiste no estudo exclusivo da regra, negligenciando o caso concreto e focando apenas no aspecto universal, demonstrando um desapego em relação ao individual - um distanciamento da realidade social (Baptista da Silva, 2004). Ao se centrar apenas no enunciado legal, perde-se de vista a singularidade de cada situação e ignorar essa particularidade significa empobrecê-la. O influxo do pensamento moderno levou o direito processual civil a uma redução conceitual alinhada às ciências exatas, alimentando o dogmatismo procedimental a ponto de afastar o processo do próprio fato concreto que lhe dá origem e sustento (Isaia, 2012). Essa opção por uma técnica impede o intérprete de visualizar como o regramento se efetiva em um tecido social vivo que é permeado por desigualdades e conflitos específicos.

O positivismo lógico, ao identificar a ciência com a linguagem, adota uma postura reducionista que concebe a linguagem como uma estrutura textual auto suficiente, e, por sua vez, a significação emerge inteiramente do sistema interno por ela criado, desconsiderando, assim, outras esferas de sentido, como a produção social dos significados, que, na realidade, precede e fundamenta o próprio significado social (Warat,

2000). Em outras palavras, essa visão seria como uma biblioteca fechada, onde todos os livros se referem uns aos outros, criando um sistema de significados que só faz sentido dentro daquele espaço.

Para os filósofos racionalistas, a tendência era abandonar a retórica forense e evitar compreender o direito como uma ciência da cultura, uma disciplina hermenêutica, em vez disso, optavam por submeter as epistemologias das ciências explicativas, quando, na verdade, a disciplina jurídica deveria ser concebida como uma ciência da compreensão - uma ciência do espírito, que compreende os produtos da criatividade humana (Baptista Da Silva, 1997). A referência recai sobre a aplicação de uma ciência matematizada e precisa em um campo de incertezas que demanda compreensão e rejeita a rigidez da lógica e exatidão. A busca pela explicabilidade, como se verá ao longo deste trabalho, assemelha-se a um redemoinho: avança e recua, retornando em diferentes formas e nuances, sem jamais perder sua essência racionalista. A era da técnica e a era do virtual são duas faces da mesma moeda (Streck, 2015).

É indispensável mencionar que, ao racionalismo, tão presente na formação da ciência jurídica moderna, em especial no direito processual civil, um dos ingredientes que culminaram nesse fenômeno foi a tripartição dos poderes independentes, pois ao Poder Judiciário foi incumbida a tarefa apenas de reprodução da lei (Baptista da Silva, 2004). Esse tipo de racionalismo pressupõe uma separação rígida entre criação e aplicação da norma, relegando o juiz à função de boca da lei, como defendia Montesquieu em sua obra intitulada “Do Espírito das Leis”.

No século XVI, Maquiavel apontava que, na França, havia três Poderes distintos: o Legislativo (Parlamento), o Executivo (rei) e um Judiciário independente (Camargo, 2021). E, conforme observa Dalmo de Abreu Dallari, o próprio Maquiavel elogia essa organização por garantir maior liberdade e segurança ao rei, e não ao povo [...] (Camargo, 2021). Então, nota-se que a separação dos poderes do Estado não foi concebida propriamente para garantir a independência entre eles, mas como um mecanismo de racionalização da administração. A tripartição funcionou mais como uma fachada de modernidade do que como uma estrutura, destinada a legitimar a manutenção de estruturas excludentes.

O sistema de carreira subordinada a uma cúpula designada pelo Executivo contribuiu para afastar o Judiciário das disputas políticas, mas, com o tempo, também o distanciou da sociedade como um todo, uma sociedade cada vez mais conflituosa e complexa, o que, em parte, explica

a crise de legitimidade que o afeta nos dias atuais. Ao longo do tempo, essa forma de organização contribuiu para a formação de uma cultura institucional autorreferente, tecnocrática e pouco sensível à pluralidade e à complexidade dos conflitos sociais contemporâneos.

São raros os que ousariam afirmar que submissão, em processo avançado de perda das liberdades, completamente expostos a formas sutis e profundas de colonização que penetram até os aspectos mais íntimos da vida, alheios à compreensão do desastre que se aproxima e, de forma apática, coniventes com a mercantilização global de existências, conduzidos por uma elite que impôs uma narrativa ilusória sob o disfarce de progresso (Sadin, 2017). Esse processo de colonização moderna não é mais caracterizado por formas de opressão visíveis, mas sim por mecanismos que agem de maneira velada.

A busca pela certeza do direito, enquanto ideal do racionalismo, culminou na formação de um sistema judiciário excessivamente burocrático, o qual vivenciamos hoje (Baptista da Silva, 1997). No entanto, essa pretensão de objetividade gera uma hipertrofia procedimental, pois o processo valoriza mais a forma do que a substância. Lembra-se que as posturas procedimentalistas não atribuem à jurisdição constitucional um papel concretizador, limitando-a à função de controle das “regras do jogo” democrático, por outro lado, as posturas substancialistas reconhecem a função concretizadora da jurisdição constitucional e consideram o Judiciário como um locus privilegiado para assegurar o fortalecimento institucional das democracias contemporâneas (Streck, 2014).

Para alguns de seus defensores, a tese da teoria constitucional voltada aos países periféricos passa justamente por uma necessária tomada de posição a favor de uma visão substancialista (Echaiz, 2008). Ao propor uma tomada de posição substancialista, defende-se que o Direito Constitucional deve ser sensível às realidades concretas vividas pelas populações desses países, visto que a mera observância procedimental das normas constitucionais não é suficiente para garantir a efetividade dos direitos fundamentais. É pertinente mencionar as teorias procedimentalistas e substancialistas, que alimentam intensos debates no Brasil em torno da decisão judicial, especialmente no que se refere à definição do papel da jurisdição constitucional e à reflexão sobre os contornos da tutela coletiva (Momolli, 2020).

Após essa abordagem memoranda, percebe-se que o processo ainda é permeado pela visão racionalista e carrega influências do direito romano.

O racionalismo e a tripartição dos poderes impactaram a formação da ciência jurídica moderna, resultando em um sistema judiciário moroso e carente de hermenêutica. Entretanto, hoje, há uma estrutura social que se baseia em redes operadas por tecnologias de comunicação e informação, como redes digitais de computadores e microeletrônica, as quais possibilitam a geração, processamento e distribuição de informações com base no conhecimento acumulado nos diversos nós dessa rede (Castells; Cardoso, 2005).

Nessa maré, durante o seminário “Impactos setoriais da inteligência artificial”, promovido pela Universidade Santo Amaro (Unisa), consolidou-se entre os especialistas o entendimento de que o impacto da inteligência artificial demanda uma revolução jurídica. André Ramos Tavares, ministro do Tribunal Superior Eleitoral, sustenta que o encolhimento das fronteiras do direito, resultado de dois fatores fundamentais: de um lado, as normas em vigor ignoram os efeitos disruptivos já consolidados na sociedade; de outro, os operadores do direito permanecem ancorados em mecanismos e instrumentos, em sua grande maioria elaborados no século XIX e herdados do direito romano (Batistella, 2024). Diante desse cenário, não basta uma simples atualização dos modelos vigentes, impõe-se uma ruptura estrutural, uma verdadeira revolução jurídica, afirmou o ministro (Batistella, 2024).

O processo civil brasileiro se depara com uma realidade ambígua, situada entre a modernidade e a hipermodernidade, experimentando uma perda de autonomia e uma certa apatia decisória diante das complexidades dos conflitos atuais (Hoffmam; Bolzan de Moraes, 2016). A perda de autonomia no processo civil, que já se manifesta na sobrecarga de formalismos, pode ser ainda mais acentuada com a introdução dos sistemas de comportamento inteligente que surgem como uma possível solução para os problemas que eles próprios ajudaram a criar. Essa situação se assemelha à ingestão de um remédio destinado a tratar uma enfermidade específica, mas que, em vez de promover a cura, acaba por desencadear uma nova condição indesejada.

Na contemporaneidade marcada pela globalização econômica, pelo modelo de Estado Mínimo de inspiração neoliberal e pela pulverização do projeto constitucional nacional, as teorias procedimentais avançam, expandindo seu espaço e conquistando adeptos por meio de justificativas diversas (Echaiz, 2008). É oferecido um conforto técnico que mascara verdadeiras intenções. Caminha-se para uma forma de testemunho integral da vida que, na realidade, configura um processo de exploração

com duas finalidades centrais: a primeira consiste na consolidação de um novo estágio do capitalismo, denominado tecnoliberalismo, cujo objetivo é ocupar todos os espaços da existência, sem deixar qualquer dimensão da vida isenta de controle (Sadin, 2017).

O tecnoliberalismo representa a fusão da lógica de mercado com o poder da tecnologia digital enquanto o neoliberalismo defendia a liberdade econômica e a mínima intervenção estatal. Com isso, a segunda função desse ambiente tecnológico se refere à organização automatizada de setores cada vez mais extensos da sociedade (Sadin, 2017). Essa nova etapa do capitalismo se vale da inteligência artificial para automatizar mecanismos de dominação e expandir a lógica mercadológica para as dimensões mais íntimas da vida privada.

O Judiciário, além das parcerias estabelecidas com universidades e dos seus próprios meios de criação de inteligência artificial, como a plataforma Sinapses, também tem contado com a colaboração de empresas privadas que desenvolvem soluções tecnológicas voltadas para o aprimoramento do processo judicial. Inclusive, uma das plataformas que recentemente ganhou expressivo destaque no meio jurídico foi a startup santa-mariense denominada Jusfy, premiada durante o *South Summit Brazil*, realizado na cidade de Porto Alegre. A referida iniciativa conquistou o prêmio na categoria “Mais Escalável”, na *Startup Competition* do evento.

A inteligência artificial tem sido incorporada ao Judiciário como instrumento técnico destinado a ampliar a capacidade decisória e a promover a digitalização dos processos, com vistas à eficiência, à celeridade e à observância da razoável duração dos procedimentos (Mundin, 2025). O direito, quando mediado por estruturas virtuais, deixa de ser apenas um sistema normativo e passa a operar também como arquitetura digital. Os magistrados passaram a ser concebidos como “gerentes do caso”, uma abordagem respaldada pelo discurso da necessidade de celeridade para a redução de custos e tempo (Nunes; Paonelli, 2021).

A definição de “metas”, tal como frequentemente empregada no país, desloca o foco da busca por uma decisão adequada ao conflito (fundamentada na participação ativa das partes) para uma lógica de discricionariedade gerencial, na qual os juízes passam a administrar os processos com base em critérios de eficiência, priorizando tempo e custo, com menor controle por parte dos envolvidos (Nunes; Paonelli, 2021). A quantidade fica em detrimento da qualidade e o conhecimento das partes sobre o funcionamento do processo judicial fica ainda mais fragilizado,

visto que o “juridiquês” se alia com a linguagem computacional devido às decisões automatizadas ou tecnicamente assistidas.

A governamentalidade algorítmica, aliada ao neoliberalismo processual, compromete o exercício pleno do devido processo legal ao suprimir a consideração das especificidades de cada lide, em nome da celeridade decisória e da aplicação automatizada de precedentes, privilegiando a eficiência em detrimento da justiça contextual e da escuta singular das partes (Mundin, 2025). O “*case*” é reduzido a um dado estatístico ou a uma variável replicável, o que suscita preocupações para quem possui leituras Dworkinianas e o mínimo de conhecimento sobre garantias constitucionais.

O uso da inteligência artificial na sistematização e aplicação de precedentes revela uma tendência antiprocessual, na medida em que privilegia a eficiência quantitativa e a produção massificada de decisões, relegando a principiologia constitucional do processo (como o contraditório, a ampla defesa e a fundamentação adequada) a um plano secundário (Mundin, 2025). A inteligência artificial, nesse modelo, torna-se instrumento de uma governamentalidade neoliberal que visa a moldar o funcionamento do Judiciário conforme os ditames do mercado, isto é, com foco em velocidade e previsibilidade.

O acanhamento do Estado Democrático de Direito, sob a ótica do constitucionalismo, revela não apenas uma redução quantitativa, mas também um empobrecimento qualitativo, com menos democracia e menos transparência, como alude Teixeira (Batistella, 2024). Parece que a ética figura como um obstáculo a ser superado, interpondo-se ao avanço representado pela inteligência artificial (Streck, 2025). Então, ainda que a presença da inteligência artificial no sistema de justiça seja inevitável e tenha seus méritos, também traz à tona desafios complexos.

3.1 Bullet Time e a virada tecnológica no poder judiciário e(m) crise

A noção de *Bullet Time*, amplamente conhecida no campo cinematográfico por sua capacidade de desacelerar o tempo para evidenciar detalhes imperceptíveis em velocidade normal (Wachowski; Wachowski, 1999), serve como metáfora para compreender o atual cenário do Poder Judiciário diante da virada tecnológica. Em um momento de crise estrutural e epistemológica, no qual o próprio conceito de justiça parece

tensionado entre a tradição e a aceleração promovida pela tecnologia, a metáfora do permite observar, em um plano ampliado, os efeitos e desafios impostos pela digitalização dos processos, pela inteligência artificial e pela automação dos serviços jurisdicionais. Se, por um lado, a revolução digital promete maior celeridade e eficiência, por outro, suscita questionamentos sobre a humanização da justiça, a desigualdade no acesso às ferramentas tecnológicas e a crescente dependência de algoritmos na tomada de decisões judiciais.

Dessa forma, urge lembrar que a tecnologia estava sendo inserida de forma gradual no Poder Judiciário brasileiro, tendo em vista que a diversidade dos atos processuais, aliada à necessária celeridade da prestação jurisdicional, encontra na inovação tecnológica do fax um instrumento de grande relevância. Sendo regulamentado seu uso pela Lei 9.800, de 26 de maio de 1999, a qual permitia às partes utilizar um sistema de transmissão de dados para a realização de atos processuais (Brasil, 1999). Assim, o *Fac-símile* se tornou um movimento que foi precursor das inovações tecnológicas subsequentes que continuariam a moldar o sistema judicial.

No transcorrer do tempo, desde o início dos anos 2000, o Brasil tem passado por mudanças estruturais na dinâmica de tramitação dos procedimentos em seu sistema público de justiça (Nunes; Paolinelli, 2021). Essas transformações começaram com a implementação do processo judicial eletrônico (PJe), intensificaram-se nos últimos anos e ganharam impulso total com a “virtualização” forçada pela pandemia de COVID-19 (Nunes; Paolinelli, 2021). Com o avanço da tecnologia e a crescente demanda por maior eficiência, em consonância com o cenário pandêmico, o qual revelou a flexibilidade e a capacidade de adaptação do Judiciário, demonstrando que o sistema é capaz de se ajustar rapidamente às novas circunstâncias. As práticas e tecnologias que foram adotadas durante a pandemia não apenas possibilitaram a manutenção das operações judiciais, mas abriram caminho para novas tecnologias e métodos de trabalho.

Na cena atual do sistema de justiça nacional, cerca de 84 milhões de processos tramitam em 91 tribunais, mais de 80% sob a jurisdição da Justiça Estadual, sendo conduzidos por aproximadamente 18 mil juízes e 275 mil servidores em todo o Brasil (CNJ, 2024). Tal panorama revela a incapacidade de uma gestão processual eficiente que corresponda à real necessidade de celeridade e efetividade. Esses números são frequentemente mencionados como sinônimos de “crise” e explosão de litígios, promovendo a ideia de que o sistema necessita de soluções urgentes (Nunes; Paonelli,

2021). A propósito, perquire-se se a crise do processo é o reflexo da crise do próprio Estado, que ainda se vê incapaz de efetivar suas promessas constitucionais (Isaia, 2013). A sobrecarga do sistema judicial não é apenas um problema interno do judiciário, mas também um sintoma de falhas mais profundas na estrutura e funcionamento do Estado.

Nesse contexto, cumpre salientar que os esforços voltados para “desafogar” o Judiciário e explorar alternativas eficazes para a resolução de conflitos não são contemporâneos, constituindo-se como uma preocupação que remonta a várias décadas. Especialmente ao final da década de 1970, com o delineamento das chamadas “ondas renovatórias” e impulsionados pelo movimento Florença de acesso à justiça, diversos países passaram a incorporar diretrizes orientadas para a busca de soluções mais ágeis, menos onerosas, simplificadas e amplamente acessíveis na resolução dos conflitos (Nunes; Paonelli, 2021).

O movimento Florença, especificamente, visava garantir que a justiça fosse acessível a todos, especialmente às camadas sociais mais desfavorecidas, que historicamente se viam à margem de um sistema judicial que, muitas vezes, se tornava um obstáculo em vez de uma solução. Nessa conjuntura, as “ondas renovatórias” representam movimentos que buscam transformar e modernizar o sistema judiciário para torná-lo mais acessível, eficiente e justo. A primeira onda renovatória envolve a assistência jurídica, a segunda envolve a representação jurídica para os interesses difusos, especialmente nas áreas de proteção ambiental e de direito do consumidor, e a terceira, denominada “enfoque de acesso à justiça”, reflete a tentativa de atacar as barreiras ao acesso à Justiça (Mendes; Silva, 2015).

Assim, a primeira onda renovatória resultou na criação de mecanismos que assegurassem a assistência jurídica gratuita aos cidadãos, garantindo que aqueles sem recursos financeiros para arcar com os custos de advogados particulares possam, ainda assim, ter direito à defesa no âmbito judicial. No entanto, a assistência judiciária gratuita se revelou insuficiente, uma vez que advogados mais experientes tendem a direcionar seu tempo e esforço para causas remuneradas, enquanto a advocacia pro bono, de maneira geral, acaba sendo realizada por profissionais menos experientes (Cappelletti; Garth; Trocker, 1976).

A segunda onda renovatória está centrada na representação jurídica dos interesses difusos, ou seja, a defesa de direitos que pertencem não a indivíduos específicos, mas a grupos ou à sociedade como um todo, como direitos ambientais e consumeristas. Em razão de que a concepção

tradicional do processo civil não contemplava a proteção dos direitos difusos, evidenciando a necessidade de uma representatividade adequada, uma vez que a ausência de participação individual nas demandas exigia a reformulação da própria noção de coisa julgada, de modo a compatibilizá-la com a tutela coletiva (Cappelletti; Garth, 2002).

A terceira onda parte do reconhecimento de que não basta garantir o direito formal de ingressar em juízo se, na prática, persistem barreiras econômicas, sociais e processuais que inviabilizam uma tutela jurisdicional efetiva. Esse enfoque impulsionou uma ampla gama de reformas, abrangendo desde a revisão dos procedimentos e a reestruturação dos tribunais até a incorporação de atores leigos ou para profissionais no sistema de justiça, bem como fomentou modificações no direito substantivo, com vistas à prevenção de litígios, e estimulou a adoção de métodos extrajudiciais de resolução de conflitos.

Em 1990, sob a influência do Consenso de Washington, o Brasil implementou reformas institucionais visando atrair investimentos e alinhar estratégias amigáveis ao mercado. No judiciário, a ênfase foi na eficiência decisória, priorizando rapidez e previsibilidade para garantir a segurança jurídica necessária à implementação da política desenvolvimentista. Começou-se a promover a criação do sistema multiportas alternativas ao sistema de justiça estatal, como a arbitragem, mediação, conciliação e a ampliação de procedimentos administrativos, também, buscou-se intensificar o uso de medidas coercitivas atípicas na execução, incentivar a adoção de práticas desjudicializadoras, especialmente por meio dos cartórios (como inventários, divórcios e usucapiões), e reforçar a consensualidade e a convencionalidade como soluções para “desafogar” o judiciário (Nunes; Paonelli, 2021).

Desde então, em todo o mundo (e não é diferente aqui), persiste a discussão sobre como as ADRs (*Alternative Dispute Resolution*) podem viabilizar ou corrigir realidades marcadas por significativos desequilíbrios de poder e marginalização (Nunes; Paonelli, 2021). Embora as ADRs ofereçam métodos menos onerosos, mais céleres e, frequentemente, mais flexíveis que o sistema judicial tradicional, em ambientes marcados por disparidades significativas em termos de recursos, conhecimento e acesso à justiça, tais alternativas podem, inadvertidamente, favorecer as partes mais poderosas. Isso ocorre porque os atores com maior expertise e capacidade de negociação acabam se beneficiando mais da estrutura, ampliando as desigualdades existentes.

Devido às diferenças de dimensão, condição jurídica e recursos disponíveis, alguns atores da sociedade têm acesso mais frequente aos tribunais (seja para ajuizar reclamações ou para se defender), enquanto outros recorrem a esses mecanismos de forma esporádica (Galanter, 2018). De modo análogo, essa dinâmica se reflete nos meios extrajudiciais onde os participantes mais favorecidos tendem a utilizar esses instrumentos de forma contínua e estratégica, em contraste com aqueles que os recorrem apenas pontualmente e talvez não logrem êxito na obtenção de um acordo justo.

Desse modo, pode-se classificar esses atores em duas categorias: os que recorrem aos tribunais apenas de forma esporádica (participantes eventuais ou PEs) e os que, de maneira habitual, se envolvem em diversas litigâncias ao longo do tempo (jogadores habituais ou JHs) (Galanter, 2018). É um exemplo claro dessa distinção que pode ser observado em disputas envolvendo grandes corporações e consumidores individuais.

Ao final de 2023, de acordo com o relatório *Justiça em Números*, estavam instalados 1.930 Cejuscs, sendo a maior parte na Justiça Estadual, com 1.724 unidades e na Justiça do Trabalho, havia 129 Cejuscs, enquanto a Justiça Federal contava com 77 unidades. Entretanto, apesar da ampla estrutura instalada, os índices de conciliação ainda enfrentam desafios para alcançar maior efetividade, pois a taxa de sentenças homologatórias, no mesmo ano, foi de apenas 12,1%. Apesar da ampla estrutura instalada, os índices de conciliação ainda enfrentam desafios para alcançar maior efetividade. Na fase de conhecimento, a conciliação alcançou 17,8%, enquanto, na fase de execução, as sentenças homologatórias corresponderam a 9,1%.

Em 2023 foram registrados 35 milhões de processos novos, o maior número da série histórica de quase 20 anos, representando um aumento de 9,4% em relação ao ano anterior (CNJ, 2024). Esse sintoma de ineficiência, que se arrasta por anos, assemelha-se a uma doença crônica não tratada corretamente, cujos sinais de agravamento se tornam cada vez mais evidentes. Então, com a busca de remédios para curar esse mal que padece o judiciário, hoje, no campo jurídico, a adoção de tecnologias disruptivas tem se tornado cada vez mais frequente (Nunes, 2019).

Nesse sentido, no processo de consolidação dessa transformação, geralmente a dividimos em três etapas: a primeira na virtualização (ou digitalização) dos procedimentos; a segunda na automação de tarefas; e a terceira na transformação dos modos de ação e trabalho (Nunes; Paonelli,

2021). Cada fase foi dando origem às próximas, conforme foram ocorrendo novas experiências digitais, entretanto, no âmbito da Justiça brasileira foi introduzida de forma gradual, tendo a experiência com a incorporação de recursos eletrônicos se iniciado já em 1996.

Basta observar que, naquela época, Ellen Graciele Northfleet, então Vice-Presidente do Tribunal Regional Federal da 4ª Região, abordou o uso do fax, alhures à publicação da lei. Considerando que, na era do terceiro milênio, os membros do Poder Judiciário devem estar capacitados para adotar novos formatos de expedição e arquivamento de dados (distintos dos antiquados cadernos processuais repletos de carimbos, certidões e assinaturas) o resguardo, ainda que almejado, não pode constituir um obstáculo à celeridade e à eficiência (Northfleet, 1996).

A posteriori, em 26 de maio de 1999, foi instituída a Lei nº 9.800 (conhecida como Lei do Fax), a qual permitiu o envio de dados processuais por meio de transmissão fotomecânica de dados e imagens, denominada fac-símile, embora as vias originais devessem ser apresentadas na vara competente, em até cinco dias após o término do prazo estabelecido. Contudo, por não dispor de assinatura eletrônica e não exigir que os órgãos do Poder Judiciário dispusessem dos instrumentos necessários para sua aplicação legítima, conforme preceitua o artigo 5º da referida lei, demonstrou-se sua reduzida influência sobre o PJe, embora tenha servido de inspiração para a incorporação da ciência digital nos tribunais.

No ano subsequente, foi proposta a viabilização da validação das documentações eletrônicas por meio da criação da Infraestrutura de Chaves Públicas Brasileira (ICP-Brasil), regulamentada pela Medida Provisória nº 2.200, de 27 de julho de 2001, a qual estabeleceu as bases legais para a certificação digital e, por conseguinte, para a assinatura digital. Na mesma época, foi publicada a Lei nº 10.259, de 12 de julho de 2001, com o objetivo de fomentar experiências digitais nos Juizados Federais, permitindo, por meio do artigo 8º, §2º, que os tribunais se organizassem para realizar intimações e receber petições de forma eletrônica, bem como possibilitando que reuniões de juízes residentes e domiciliados em cidades diversas fossem conduzidas eletronicamente, conforme disposto no artigo 14, §3º.

Em 2001 a Associação dos Juízes Federais (AJUFE) submeteu à Câmara dos Deputados um anteprojeto de lei, identificado pelo nº 5828, de 04 de dezembro de 2001, que versava sobre a informatização processual com vistas a combater a morosidade processual, dentre outras questões.

Este anteprojeto tramitou no Poder Legislativo por cerca de cinco anos, culminando na promulgação da Lei nº 11.419, de 2006, Lei do Processo Judicial eletrônico (Pje). Não obstante, foi com a eclosão da pandemia de Covid-19 (SARS-CoV-2) e as rigorosas medidas sanitárias que o Poder Judiciário se viu compelido a estabelecer um novo padrão de atividades jurisdicionais. Logo, a aceleração da informatização processual se tornou uma necessidade premente.

Os *lockdowns* impostos para conter a disseminação do vírus, que o mundo foi forçado a uma readaptação, não apenas para dar continuidade às suas atividades, mas também para estabelecer um novo padrão de convivência e funcionamento social. Dessa forma, ao “novo normal” se pode atribuir a concepção de que a pandemia viral trouxe consigo a necessidade de ressignificar a turbulência e a ruptura de antigos hábitos, transformando-as em forças motrizes para a evolução e a reconfiguração da sociedade (Pedron; Torres; Menezes, 2022).

A Resolução CNJ nº 420, de 29 de setembro de 2021, definiu um cronograma para a digitalização do acervo processual físico de todos os órgãos do Poder Judiciário, garantindo a tramitação dos processos em sistemas eletrônicos. Em acréscimo, a norma vedou o ingresso de novos casos a partir de março de 2022 e, conforme o disposto no art. 3º, IV, os tribunais que, em 30 de setembro de 2021, tinham um acervo físico superior a 40% deveriam concluir a digitalização até 31 de dezembro de 2025.

De acordo com dados oficiais do Conselho Nacional de Justiça em 2019 apenas 10% dos novos processos foram ingressados fisicamente e ao final de 2023, 90,6% dos processos em tramitação eram eletrônicos, abrangendo 92% no segundo grau, 90,4% no primeiro grau e 100% nos Tribunais Superiores. A propósito, o mesmo relatório aponta que a Justiça Eleitoral e a Justiça do Trabalho se destacam por possuírem diversos tribunais com 100% de processos eletrônicos em ambas as instâncias, tanto no primeiro quanto no segundo grau. Então, no que tange à primeira etapa rumo à consolidação de um cenário de Justiça Digital, o Brasil já se encontra em estágio bastante avançado (Nunes; Paonelli, 2021). Logo, observa-se que o passo inicial consistiu na transformação dos processos físicos em virtuais, abrangendo tanto a digitalização dos processos já existentes quanto a inauguração de novos processos exclusivamente no formato eletrônico.

Entretanto, segundo os registros do CNJ o ano de 2023 findou com um acervo de 83,8 milhões de processos em tramitação, incluindo os suspensos, sobrestados e em arquivamento provisório, evidenciando um aumento de 03 milhões de novos casos em relação a 2022. Com isso, na fase de automatização, observa-se que a maioria das ferramentas implementadas ou em desenvolvimento no Judiciário brasileiro visa combater o expressivo volume de demandas judiciais (Nunes; Paonelli, 2021). O Programa Justiça 4.0: Inovação e Efetividade na Realização da Justiça tem como objetivo ampliar o acesso à Justiça por meio de iniciativas e projetos que incentivam o uso colaborativo de soluções baseadas em novas tecnologias e inteligência artificial.

Dentre suas principais iniciativas, destaca-se o Domicílio Eletrônico, concebido para centralizar as comunicações processuais virtuais destinadas a pessoas físicas e jurídicas, bem como a Plataforma Digital do Poder Judiciário (PDPJ-Br) que fomenta o desenvolvimento colaborativo de sistemas. Paralelamente, o DataJud consolida estatísticas e aprimora a transparência no âmbito judiciário. Ainda, a modernização do sistema também abarca o Núcleo de Justiça 4.0, o Juízo 100% Digital, o Balcão Virtual e os Pontos de Inclusão Digital, iniciativas que promovem maior acessibilidade e eficiência.

E, por sua vez, a Plataforma Codex emprega Inteligência Artificial (IA) na captura e processamento de peças processuais, potencializando a automação e a celeridade dos trâmites judiciais. Enquanto, segundo o CNJ, o Sinapses é uma plataforma voltada para o desenvolvimento e a ampla disponibilização de modelos de IA e por meio dela o processo de entrega dos modelos é significativamente acelerado em comparação com os sistemas tradicionais (Nunes; Paonelli, 2021). Ao passo que o objetivo é fornecer uma gama de modelos para utilização no PJe, permitindo que sejam compatíveis com diversas versões e possibilitando que cada tribunal desenvolva seus próprios modelos, compartilhe-os e utilize os modelos criados por outros tribunais (Nunes; Paonelli, 2021).

Ao viabilizar essa dinâmica descentralizada, vê-se que o Judiciário pretende fomentar um intercâmbio tecnológico com o intuito de otimizar a gestão processual através de um ecossistema de inovação contínua. Por sinal, o CNJ tem se referido ao Sinapses como uma verdadeira “fábrica de modelos de IA”, sendo essa iniciativa vista como um possível elemento central em uma estratégia de governança da inteligência artificial (Nunes; Paonelli, 2021). O peso dessa expressão enfatiza o potencial de transformação do

sistema tradicional, inaugurando uma nova era no Judiciário, marcando um divisor de águas na forma como a justiça é concebida e aplicada.

O Inova PJe, por seu turno, é o Laboratório de Inovação do Processo Judicial Eletrônico que opera em conjunto com o Sinapses e está concebido para atuar como um acelerador da inteligência artificial no PJe (Nunes; Paonelli, 2021). Ele é responsável pela estruturação da base nacional de dados jurídicos destinada ao treinamento das ferramentas e conta com um Centro de Inteligência Artificial, que reúne expertise técnica e intelectual em IA para orientar o desenvolvimento de soluções e promover comunidades dedicadas ao compartilhamento de informações (Nunes; Paonelli, 2021). À vista disso, vê-se um epicentro onde a criatividade e a tecnologia se encontram, moldando um novo futuro para o sistema judicial.

Desde 2020, o CNJ vem elaborando um relatório de pesquisa que tem por escopo mapear os projetos de inteligência artificial em andamento nos tribunais e acompanhar sua evolução, em estrita consonância com as diretrizes estabelecidas pela Resolução nº 332/2020 e pela Portaria nº 271/2020, ambas emanadas do próprio CNJ. No mapeamento realizado em 2023, foram identificados 140 projetos de IA, desenvolvidos ou em desenvolvimento, nos tribunais e conselhos, o que representa um crescimento de 26% em relação a 2022.

Os projetos de inteligência artificial estão distribuídos por ramo da Justiça da seguinte forma: os tribunais estaduais concentram 68 projetos, representando 48,6% do total; os tribunais eleitorais desenvolvem 23 iniciativas (16,4%); a Justiça do Trabalho conta com 20 projetos (14,3%); os tribunais federais implementam 14 projetos (10%); os tribunais superiores somam 13 projetos (9,3%); e, por fim, os conselhos jurídicos abrigam 2 projetos, equivalentes a 1,4% do total.

As redes neurais figuram como os algoritmos mais citados, presentes em 79 projetos (56,4% do total), evidenciando sua versatilidade e eficácia no processamento de linguagem natural, atributos indispensáveis para a gestão do vasto volume de dados textuais manejados pelo Judiciário brasileiro (CNJ, 2024). Então, são utilizadas técnicas para categorizar automaticamente diferentes tipos de peças processuais e para reconhecimento de entidades nomeadas (NER) que identificam e extraem informações essenciais, como nomes, datas e locais. Essas abordagens tornam o acesso à informação mais eficiente, contribuindo para a celeridade e a organização dos procedimentos no Judiciário.

As árvores de decisão e florestas aleatórias também figuram entre os algoritmos frequentemente mencionados, cada um com 31 citações (presente em 21,1% dos projetos) que há inclinação por modelos robustos e interpretáveis, capazes de proporcionar maior transparência na tomada de decisões (CNJ, 2024). A interpretabilidade do primeiro modelo permite que os usuários compreendam como determinadas variáveis influenciam os resultados, dado que conforme explicado no primeiro capítulo, nas árvores de decisão o algoritmo constrói uma estrutura em forma de árvore, na qual cada nó interno realiza um teste que direciona o fluxo para um ramo específico.

Esse percurso até a decisão final oferece uma visão detalhada de como cada atributo contribui para a classificação, facilitando tanto a interpretação quanto a identificação de eventuais falhas ou vieses no modelo (modelo transparente). Em contrapartida, na floresta aleatória (imagine mais de uma árvore trabalhando em conjunto) a agregação de diversas árvores dificulta a compreensão direta do processo decisório, exigindo um esforço adicional para explicar como as decisões foram alcançadas (modelo opaco). Isso também se aplica às máquinas de vetores de suporte (SVMs) e ao Perceptron multicamada.

As máquinas de vetores de suporte foram citadas 16 vezes (presentes em 11,4% dos projetos), e técnicas de clusterização, como *K-means*, com 03 menções (2,1%), e DBScan, com 2 menções (1,4%), são destacadas como ferramentas para categorização e identificação de padrões e refletem o interesse em explorar relações e agrupamentos dentro de conjuntos de dados complexos (CNJ, 2024). As SVMs (aprendizado supervisionado) se destacam por sua capacidade de encontrar o melhor hiperplano que separa diferentes classes, enquanto as técnicas de aglomeração (aprendizado não supervisionado) agrupam dados com base em similaridades, sem a necessidade de rótulos com definição prévia.

O aprendizado supervisionado foi a abordagem mais mencionada, com 87 citações (presente em 62,1% dos projetos), indicando que grande parte das iniciativas de IA no Judiciário se concentra em modelos que aprendem a partir de dados rotulados, como na classificação de documentos e na previsão de desfechos judiciais (CNJ, 2024). Ao utilizar exemplos previamente categorizados, esses algoritmos são capazes de identificar padrões específicos, resultando em conclusões mais precisas e confiáveis. Já o aprendizado não supervisionado, citado 38 vezes (27,1% dos projetos), é empregado na identificação de padrões e agrupamentos em dados

sem rótulos pré-definidos, demonstrando sua utilidade na investigação exploratória de informações judiciais (CNJ, 2024). Com essa técnica os algoritmos conseguem, por exemplo, segmentar casos semelhantes.

Além disso, a presença de técnicas mais avançadas, como aprendizado por reforço, com 9 menções (6,4% dos projetos), e aprendizado por transferência, com 2 menções (1,4%), sugere uma busca por soluções mais sofisticadas, voltadas à otimização de processos judiciais e à adaptação de modelos de IA a cenários jurídicos específicos (CNJ, 2024). O aprendizado por reforço permite treinar agentes que tomam decisões com base em feedback contínuo, enquanto o aprendizado por transferência aproveita conhecimentos já adquiridos em outros contextos para aprimorar a aplicação em casos jurídicos.

A mesma pesquisa do CNJ menciona que outras aplicações, como a geração de texto, a sumarização, o reconhecimento de imagem, os *chatbots* conversacionais e os sistemas de resposta a perguntas (cada uma abordando aspectos distintos do trabalho jurídico) evidenciam a capacidade da IA de se adaptar aos diversos processos do Judiciário. O Índice de Produtividade da Magistratura Brasileira (IPM) registrou um aumento de 6,8% em 2023, ultrapassando a marca de 2.000 processos concluídos, em média, por cada magistrado e o Índice de Produtividade por Servidor (IPS-Jud) cresceu 5% em 2023, atingindo uma média anual de 170 processos finalizados por cada servidor ou servidora da área judiciária. (CNJ, 2024). Consequentemente, o CNJ destaca que, no Judiciário brasileiro, as atividades de IA assumem um papel essencial, pois a automação e a otimização proporcionadas contribuem para o aumento da produtividade dos profissionais, refletindo diretamente na redução do acervo processual e na maior celeridade dos trâmites legais.

No seu relatório anual, o órgão ainda destaca como aspecto positivo a integração seletiva dos modelos de linguagem ampla (LLMs) nas rotinas judiciais, mas apesar do crescimento de sua popularidade, esses modelos são utilizados de maneira criteriosa, evitando sua aplicação indiscriminada nos tribunais. Esse interesse crescente surge da habilidade dos LLMs em replicar processos de raciocínio e pensamento humano, possibilitando que desempenhem funções antes exclusivas de especialistas humanos, ao mesmo tempo representam uma solução econômica e escalonável para atender à demanda em expansão (Jiawei. et al, 2025).

Com os avanços em paradigmas de treinamento, como o aprendizado por reforço a partir do *feedback* humano (RLHF), os LLMs

passaram a se alinhar cada vez mais aos valores humanos e aos processos de raciocínio (Jiawei. et al, 2025). Ou seja, esse avanço possibilitou que os LLMs não apenas gerassem textos mais coerentes e úteis, mas também simulassem o julgamento humano ao analisar entradas baseadas em regras. Desse modo, a habilidade dos LLMs de simular a tomada de decisão humana e apurar entradas com base em regras predefinidas abriu espaço para o conceito de “LLM-como-um-Juiz” (Jiawei. et al, 2025). Não obstante, sua imparcialidade ainda é questionável, pois podem internalizar visões apresentadas nos dados de treinamento.

Em um recente passo evolutivo, a tecnologia ampliou as possibilidades e redefiniu os horizontes para a resolução de conflitos (Marques, 2019). Posto que no âmbito da terceira e última etapa da virada tecnológica, a Resolução nº 358 do CNJ, publicada em 02 de dezembro de 2020, estabeleceu diretrizes para a implementação de inovações tecnológicas voltadas à solução de controvérsias no Poder Judiciário, valendo-se dos mecanismos de conciliação e mediação para conferir maior eficiência e celeridade aos processos. A referida resolução orienta os tribunais brasileiros no desenvolvimento de sistemas que viabilizem o cadastro eletrônico das partes e seus representantes, a integração com o Cadastro Nacional de Mediadores e Conciliadores do CNJ (Conciliajud), o registro de casos extrajudiciais, a geração automatizada de atas e termos de acordo, além da integração modular com o sistema de processo eletrônico adotado pelo tribunal, permitindo a comunicação entre o SIREC e o sistema de julgamento da instituição (Nunes; Paonelli, 2021).

Nesse contexto, ganharam destaque os chamados “*Online Dispute Resolution*” (ODR), que basicamente representam a adaptação dos métodos apropriados de solução de conflitos para plataformas online (Lima, 2019). Distanciando-se dos métodos tradicionais de Alternative Dispute Resolution (ADR), previstos no Código de Processo Civil de 2025, como a mediação e a arbitragem presenciais, os ODRs empregam a tecnologia para viabilizar um ambiente digital no qual as partes podem solucionar suas disputas sem a exigência de presença física. Sendo, de fato, a união entre a tecnologia da informação e os métodos alternativos de resolução de controvérsias.

Diante desse quadro, um dos casos de grande sucesso e reconhecimento mundial é o Modria (Faria, s.d). Inicialmente, desenvolvido a partir da experiência do *eBay* e do *PayPal*, que necessitavam de soluções rápidas e eficazes para resolver milhões de conflitos entre usuários, o

sistema expandiu-se para diversas áreas do direito, demonstrando sua versatilidade e eficiência. A plataforma agiliza a tramitação processual até a resolução final do conflito, utilizando processamento de linguagem natural para diagnosticar o problema a partir da leitura de texto e com base nisso viabiliza a negociação on-line entre as partes, possibilita o acesso a um mediador humano, quando necessário, e, posteriormente, encaminha o caso para avaliação do resultado (Faria, s.d).

Nestes ambientes digitais, inserem-se elementos fundamentais que sustentam os métodos de resolução de disputas online, como: i) formas de comunicação síncronas ou assíncronas; ii) eliminação das limitações territoriais; iii) a atuação da tecnologia como uma “quarta parte” no processo; iv) diminuição de custos e maximização da eficiência temporal; e v) estratégias preventivas de resolução de conflitos (Nunes; Paonelli, 2021). De forma mais precisa, essas plataformas conferem maior flexibilidade à interação entre as partes, viabilizando tanto a comunicação instantânea quanto as respostas mais elaboradas ao longo do tempo. Além de que suprimem as restrições geográficas inerentes aos métodos tradicionais e promovem a otimização de custos.

A revolução tecnológica no sistema de justiça brasileiro representa uma profunda reconfiguração das formas tradicionais de trabalho, impulsionada por inovações contínuas que vão além da mera digitalização processual. Cada etapa desse avanço não apenas buscou a eficiência operacional do Judiciário, mas também redefiniu as relações entre os atores do processo e ressignificou o próprio conceito de prestação jurisdicional, na medida em que o exercício e o acesso ao direito passaram a ser, em muitos casos, mediados por inteligências artificiais que executam tarefas antes restritas aos juristas. Com isso, florescem parcerias entre órgãos do Poder Judiciário, universidades e empresas privadas para fomentar soluções tecnológicas e jurídicas que respondem efetivamente às demandas contemporâneas.

3.2 Metacortex, automação processual e estratégias de modernização e eficiência operacional: o exemplo privilegiado do IRDR, sistema de precedentes

A empresa onde Neo atuava como programador se chamava “Metacortex”, um nome que evoca a ambição da ciência moderna de tentar replicar a complexidade da mente humana em máquinas. Tendo em vista

que o termo “meta” transmite a ideia de algo a ser alcançado, enquanto “córtex” se refere à camada externa do cérebro, responsável pela cognição, percepção e consciência. E, por sinal, o córtex pré-frontal ventromedial está envolvido no raciocínio social e na tomada de decisões; a amígdala, na avaliação social de expressões faciais; o córtex somatosensorial direito, na empatia e na simulação; e a ínsula, na resposta autonômica (Butman; Allegri, 2001).

Essa busca por replicar e até mesmo superar as capacidades humanas por meio da tecnologia não se limita à ficção científica, dado que conforme explicado no capítulo anterior a virada tecnológica proporcionou ferramentas robóticas para auxiliar os magistrados e servidores do judiciário em suas demandas administrativas. Nesse trilha, os tribunais estabelecem parcerias entre si e com instituições acadêmicas para desenvolver e aprimorar projetos de inteligência artificial (CNJ, 2024).

Cumprir destacar a relevância da Resolução nº 335, de 29 de setembro de 2020, que institui a Plataforma Digital do Poder Judiciário e implementa a política pública voltada para a governança e gestão dos processos judiciais eletrônicos, essa medida, fruto de um esforço colaborativo que envolve todos os tribunais do país, visa à adoção das melhores soluções tecnológicas para o benefício comum (Ribeiro, 2023). A robusta articulação entre as diversas unidades judiciárias evidencia um comprometimento inequívoco com a inovação, promovendo a integração de práticas modernas que potencializam a eficiência dos serviços prestados.

A expertise acadêmica e a capacidade investigativa das instituições de ensino impulsionam o desenvolvimento de novas metodologias, protótipos e sistemas inovadores, que podem ser incorporados ao cotidiano do Judiciário. Em contrapartida, esse oferece um campo prático para a aplicação dessas tecnologias, apresentando desafios reais que orientam as pesquisas e garantem que as inovações atendam de maneira eficaz às necessidades do setor. Adicionalmente, a colaboração com entidades jurídicas, como o próprio CNJ, bem como com tribunais eleitorais e do trabalho de diferentes regiões, evidencia o compromisso em promover sinergia e compartilhar as melhores práticas dentro do sistema judiciário (CNJ, 2024).

Em 2023, 23 (16,4% do total) projetos foram realizados em colaboração com universidades e 39 (27,9% do total) foram realizados em colaboração com outras entidades da área jurídica, como tribunais, órgãos ou quaisquer outras instituições jurídicas (CNJ, 2024). Conforme dados

dessa pesquisa do CNJ, 66% dos tribunais brasileiros possuem projetos de inteligência artificial em desenvolvimento e atualmente, há 147 sistemas de IA comunicados ao Sistema Sinapses. A seguir, destacam-se alguns desses robôs, localizados por meio de uma pesquisa realizada em cada tribunal e é possível observar a disseminação de robôs por todo o Judiciário:

Quadro 02: Compilado de alguns robôs sistemas robóticos em operação no Judiciário nacional e suas funcionalidades.

TRIBUNAL/ÓRGÃO	ROBÔS E FUNÇÕES
Supremo Tribunal Federal (STF)	Robô Victor: Classificação de processos e identificação de repercussão geral. Robô Rafa: Padronização da classificação dos processos. Robô VitóriaIA: Agrupamento e classificação de processos. Robô MarIA: IA generativa para resumos, relatórios e avaliação inicial de processos.
Superior Tribunal de Justiça (STJ)	Robô Athos: Identifica documentos diferentes que tratam do mesmo assunto.
Supremo Tribunal Militar (STM)	Robô Lara: Geração de textos simplificados.
Tribunal Superior Eleitoral do Espírito Santo (TSE-ES)	Robô Bel: Esclarecimento de dúvidas eleitorais sobre situação eleitoral, local de votação e voto em trânsito.
Tribunal Superior do Trabalho (TST)	Robô Bem-te-vi: Filtros para verificação de processos por temas e tempo de entrada.
Tribunal Regional do Trabalho da 3ª Região (MG)	Robô Diana: Diagnóstico do nível de acessibilidade de portais eletrônicos.
Tribunal Regional do Trabalho da 4ª Região (RS)	Robô Copas TRT4: Triagem de peças processuais para investigação. Robô Gael: Certificação automática de dados de alvarás e notificação de pagamentos. Robô Prea: Automatiza autuação de processos no PJe para precatórios e RPVs. Robô Prea 2.0: Aperfeiçoamento do Prea, com extração e pré-autuação de dados. Robô Galileu: Extração de dados do PJe para geração de relatórios e apoio na confecção de acórdãos e sentenças.

TRIBUNAL/ÓRGÃO	ROBÔS E FUNÇÕES
Tribunal Regional do Trabalho da 5ª Região (BA)	Robô Ágata: Automatiza o envio de processos aptos ao TST.
Tribunal Regional do Trabalho da 8ª Região (PA/AP)	<p>Robô Leão: Automação da confecção e assinatura de certidões processuais.</p> <p>Robô Pipa: Filtragem de processos aguardando julgamento e intimação das partes.</p> <p>Robô Paco: Identificação de processos aptos para publicação de decisões e notificação das partes.</p>
Tribunal Regional do Trabalho da 9ª Região (PR)	Robô RJ 21 (Projeto Solaria): Plataforma para criação de fluxos automatizados para movimentação de processos.
Tribunal de Justiça do Estado do Rio Grande do Sul (TJRS)	<p>Robô Módulo Júri: Automatiza a consulta de endereços, facilitando a localização das partes pelos servidores.</p> <p>Sistema de Audiências Inteligentes: Transcrição em tempo real de depoimentos, permitindo a verificação instantânea pelos advogados, Defensoria Pública e Ministério Público.</p> <p>Gaia Minuta: Auxilia magistrados na elaboração de minutas de decisões, gerando um texto inicial automaticamente a partir dos elementos selecionados por um comando simples.</p> <p>Gaia Assistente: Objetiva acelerar a prestação jurisdicional, auxiliando magistrados e servidores no estudo, organização e interpretação das informações dos autos nos processos de 1º e 2º graus.</p> <p>Gaia Explica Aí, Tchê: Traduz sentenças e decisões para uma linguagem acessível, promovendo a compreensão ampla das decisões judiciais.</p> <p>Gaia SALUS: Busca automatizar a validação dos documentos apresentados para comprovação de despesas com medicamentos.</p> <p>Gaia SEI: Promove a agilidade na tramitação de expedientes e o aprimoramento das rotinas administrativas.</p>

TRIBUNAL/ÓRGÃO	ROBÔS E FUNÇÕES
Tribunal de Justiça do Estado de Santa Catarina (TJSC)	<p>Robô Auxiliar: Classificação de petições e automatização da inserção de minutas.</p> <p>Robô Busca de Ativos Judiciais: Pesquisa de ativos de devedores em processos.</p> <p>Robô Subscritor: Auxílio na confecção e assinatura de documentos.</p> <p>Robô Arquivamento: Automação do arquivamento de processos.</p> <p>Robô Consulta de Endereços: Pesquisa de endereços para processos.</p> <p>Robô Renajud: Localização de veículos vinculados a processos.</p>
Tribunal de Justiça do Estado do Paraná (TJPR)	<p>Robô Larry: Agrupamento de processos por similaridade.</p> <p>Robô Larry Assessor 2022: Identificação de recursos semelhantes com PLN (Processamento de Linguagem Natural).</p>
Tribunal de Justiça do Estado de São Paulo (TJSP)	Robô Judi: Assistente virtual para esclarecimento de dúvidas sobre Juizados Especiais Cíveis.
Tribunal de Justiça do Estado do Rio de Janeiro (TJRJ)	Robô Assis: IA generativa para construção de minutas de sentenças, decisões e relatórios.
Ministério Público do Rio de Janeiro (MPRJ)	Robô de Registro de Candidatura: Identificação de candidaturas irregulares e verificação de elegibilidade.
Tribunal de Justiça do Estado de Minas Gerais (TJMG)	Robô Sofia: Conversação com as partes do processo traduzindo termos jurídicos para linguagem simples.
Tribunal de Justiça do Estado da Bahia (TJBA)	<p>Robô Oxóssi: Agrupamento de peças processuais por similaridade.</p> <p>Robô Themis: Redução de processos de execução fiscal e recuperação de créditos.</p> <p>Robô Haia: Automação de minutas de Execução Fiscal no PJe.</p>
Tribunal de Justiça do Estado de Sergipe (TJSE)	Robô Taise: Automatização de movimentações processuais nos sistemas SCPV e SEEU.
Tribunal de Justiça do Estado de Alagoas (TJAL)	<p>Robô Aslan: Investigação e agrupamento de petições intermediárias.</p> <p>Hércules: Classificação e categorização de petições na execução fiscal municipal.</p>

TRIBUNAL/ÓRGÃO	ROBÔS E FUNÇÕES
Tribunal de Justiça do Estado de Pernambuco (TJPE)	Robô Expedito: Automatização de atos cartorários criminais.
Tribunal de Justiça do Estado da Paraíba (TJPB)	Robô LitisControl: Controle da litigância abusiva e marcação automática de processos semelhantes. Robô SAPP: Acompanhamento de prisões preventivas e prazos legais.
Tribunal de Justiça do Estado do Ceará (TJCE)	Robô Alice: Monitoramento de editais de licitação. Robô Talia: Transcrição automática de áudios para texto. Robô Midas: Agrupamento automatizado de atos processuais similares.
Tribunal de Justiça do Estado do Piauí (TJPI)	Robô RIC: Identificação de óbitos de partes envolvidas em processos.
Tribunal de Justiça do Estado do Maranhão (TJMA)	Robô Mário Lúcio: Automação de processos judiciais. Robô Nirie: Detecção de falta de documentos em processos. Robô Maria Firmina: Investigação automática de petições iniciais e identificação de precedentes qualificados.
Tribunal de Justiça do Estado do Amapá (TJAP)	Robô Tucujuris: Identificação de demandas repetitivas para agrupamento automático. Chatbot Juci: IA para respostas automatizadas sobre processos.
Tribunal de Justiça do Estado do Rio Grande do Norte (TJRN)	Robô Poti: Penhora online automática de valores em execuções fiscais. Robô Clara: Automação da elaboração de despachos iniciais. Robô Jerimum: Classificação e rotulagem de processos. Robô Soseverino : Automatização da retirada de suspensão de processos.
Tribunal de Justiça do Estado do Pará (TJPA)	Robô Migrador 4.0: Conversão e migração automatizada de processos físicos para eletrônicos.
Tribunal de Justiça do Estado de Roraima (TJRR)	Robô Ju: Primeiro robô humanoide do Brasil para simulação de atendimento humano. Robô Arandu: Identificação de demandas repetitivas e predatórias.
Tribunal de Justiça do Estado do Amazonas (TJAM)	Robô TIC: Identificação de litígios predatórios.

TRIBUNAL/ÓRGÃO	ROBÔS E FUNÇÕES
Tribunal de Justiça do Estado de Rondônia (TJRO)	Robô Sinapses: Investigação de peças processuais e identificação de precedentes.
Tribunal de Justiça do Estado do Acre (TJAC)	Robô Automação Massiva: Elaboração automatizada de minutas de bloqueio de valores.
Tribunal de Justiça do Estado do Mato Grosso (TJMT)	Robô #TáPago: Automatização de pagamento de processos. Robô Illustris_IA: Agilização da baixa de processos com trânsito em julgado. Robô Veredictus: Intimação automática em acórdãos no PJe.
Tribunal de Justiça do Estado de Goiás (TJGO)	Robô Berna: Identificação de processos com dados semelhantes.
Tribunal de Justiça do Estado do Mato Grosso do Sul (TJMS)	Robô JU/M5 (Renajud): Extração automatizada de dados do SAJ. Robô JU/M9: Consulta de endereços em múltiplos sistemas.
Tribunal de Justiça do Distrito Federal e dos Territórios (TJDFT)	Robô Artiu: Direcionamento automatizado de mandados. Robô Toth: IA para recomendação de classes e assuntos processuais. Robô Luzia: Primeira advogada robô do Brasil para investigação de processos.

Fonte: Autora (2024)

A escolha de algumas nomenclaturas desses sistemas operacionais transcende a mera designação técnica, pois reflete a grandiosidade da cultura brasileira, referências históricas e simbólicas. Ao atentar que o Sistema Oxóssi evoca a força e a espiritualidade do orixá ligado à fartura, à coragem e à cura, amplamente venerado nas religiões de matriz africana e sincretizado com São Sebastião. Já, o Robô Jerimum faz alusão às regiões Norte e Nordeste, reforçando a identidade cultural dessas localidades e a Robô Maria Firmina celebra a memória de Maria Firmina dos Reis, a primeira romancista negra do Brasil, cuja obra e trajetória foram marcadas pela luta por justiça social e igualdade.

O Robô Arandu carrega um significado de origem tupi, em que “Arandu” remete ao entendimento, conhecimento e sabedoria, qualidades que refletem a própria capacidade do sistema de aprender e evoluir. Da mesma forma, o Robô Toth, inspirado na mitologia egípcia, remete ao deus

da sabedoria, da escrita e do conhecimento, simbolizando a busca incessante pelo saber e pelo aprimoramento intelectual. Assim, essas designações não apenas nomeiam, mas também incorporam valores e heranças culturais diversas, enriquecendo o significado dessas criações tecnológicas.

No mês de julho do corrente ano, o Tribunal de Justiça do Estado do Rio Grande do Sul realizou o evento intitulado “Conexão GAIA: Inovações de Inteligência Artificial e Entregas do Judiciário Gaúcho”, ocasião em que foram publicamente anunciadas novas iniciativas tecnológicas voltadas à modernização da Justiça e ao incremento da eficiência na prestação jurisdicional. Dentre as inovações apresentadas, destaca-se o lançamento de oito agentes baseados em Inteligência Artificial, concebidos com o propósito de conferir maior celeridade e racionalidade aos trâmites processuais. Desses, cinco já se encontram em operação, sob a denominação GAIAS (Gestão Avançada de Inteligência Artificial), representando um significativo avanço no processo de transformação digital do Judiciário sul-rio-grandense.

Durante a solenidade, o Desembargador Antônio Vinicius Amaro da Silveira enfatizou o caráter instrumental das inovações tecnológicas apresentadas, ao afirmar que tais ferramentas “serão instrumentos de auxílio às atividades realizadas, que jamais irão substituir o ser humano, que continuará comandando a situação”. Ao proferir tais palavras, o magistrado não apenas reafirmou a centralidade da atuação humana na dinâmica jurisdicional, como também sublinhou a imprescindibilidade da capacitação contínua de todos os operadores do sistema de justiça, condição essencial para a adequada incorporação dessas inovações ao cotidiano forense. E em apenas um mês desde seu lançamento, a inteligência artificial do Tribunal de Justiça do Rio Grande do Sul gerou 95 mil minutos, que, após revisão humana, vêm sendo convertidas em decisões judiciais (Jacobsen, 2025).

As aplicações da inteligência artificial abrangem tanto as atividades-meio (administrativas) quanto às atividades-fim no âmbito judicial, destacando-se as seguintes frentes: automação de tarefas repetitivas (Aumento da eficiência operacional e otimização do tempo dos servidores); suporte à decisão e celeridade processual (Auxílio na tomada de decisões e redução do tempo de tramitação dos processos); aprimoramento dos serviços de atendimento (Implementação de IA em balcões virtuais, chatbots e tradução do “juridiquês” para uma linguagem acessível); otimização de processos administrativos (Aplicação da IA em licitações,

planejamento estratégico e resolução de questões administrativas); análise aprofundada de documentos e precedentes (Aperfeiçoamento da análise jurídica, identificação de litispendência e detecção de demandas predatórias); apoio à tomada de decisão judicial (Assistência a magistrados na elaboração de minutas e na fundamentação de julgamentos); e foco em análises processuais qualificadas (Contribuições diretas para a atividade judicante, aprimorando a interpretação e aplicação do direito) (CNJ, 2024). A adoção de robôs e algoritmos, desde a classificação de processos até a geração de relatórios e minutas, reflete não somente um aprimoramento dos mecanismos internos de gestão, mas também a necessidade de readequação das instituições frente a uma sociedade cada vez mais dinâmica e complexa.

A sociedade estruturada sob a égide do Direito consolidou a institucionalização do poder, incumbindo ao ordenamento jurídico a mediação entre as dinâmicas políticas e os valores éticos por meio de uma tríade conceitual – Fato, Valor e Norma (Pinheiro, 2021). Sob a perspectiva kelseniana há uma tentativa de articular a complexidade do fenômeno jurídico em três dimensões com o desígnio de o Direito “deve se(r)” concentrar na estrutura e na validade das regras, independentemente dos elementos empíricos e dos valores morais que podem inspirar sua criação. Entretanto, o Direito Digital, apesar de se fundamentar nos mesmos pilares tradicionais, acrescenta um novo componente: o Tempo (Pinheiro, 2021). Tendo em consideração a velocidade da inovação tecnológica e a efemeridade das interações virtuais exigem uma resposta jurídica mais ágil e contínua.

A convergência entre sistemas simbólicos de representação do conhecimento jurídico e ferramentas estatísticas de aprendizado de máquina configura uma das fronteiras mais avançadas da pesquisa acadêmica e das empresas de tecnologia, ao lado da exploração de aplicações de Inteligência Artificial Generativa (IAGs) e de Grandes Modelos de Linguagem (*Large Language Models* – LLMs) (CNJ, 2024). Enquanto os sistemas simbólicos se baseiam em regras lógicas e estruturas formais que buscam modelar o raciocínio jurídico, as ferramentas estatísticas operam com base na verificação massiva de dados, identificando padrões e inferindo probabilidades.

Nesse contexto, a Inteligência Artificial Generativa (IAG) e os Grandes Modelos de Linguagem (LLMs) despontam como ferramentas fundamentais para a evolução da prática jurídica. As IAGs possibilitam a

geração de conteúdos jurídicos personalizados, como minutas de petições e pareceres, enquanto os LLMs, treinados em vastos repositórios normativos e jurisprudenciais, viabilizam uma análise textual mais profunda e contextualizada.

Curiosamente, no dia 19 de março de 2025, a empresa Jusbrasil, detentora da maior base de dados jurídicos do Brasil, lançou a plataforma Jus IA que foi concebida com fundamento em Inteligência Artificial Generativa com a proposta de conferir maior celeridade, confiabilidade e acessibilidade às pesquisas jurídicas realizadas por advogados. A ferramenta concede ao usuário a possibilidade de descrever os fatos de um caso, permitindo que a IA elabore peças jurídicas completas, meticulosamente fundamentadas em normas, jurisprudências e precedentes oficiais, de forma semelhante ao ato de formular um desejo diante da lâmpada de Aladim. Em adição, viabiliza a formulação de indagações jurídicas e a verificação das referências normativas empregadas, garantindo que os fundamentos estejam plenamente alinhados ao ordenamento vigente, conferindo maior segurança e precisão à construção argumentativa.

Sob essa ótica, a parceria com empresas líderes no setor tecnológico, como *Microsoft*, Google e AWS, juntamente com consultorias especializadas, como Softplan e Stefanini, enfatizam o intento de incorporar os avanços tecnológicos do setor privado aos órgãos do Poder Judiciário, fomentando um ambiente de inovação e melhoria contínua (CNJ, 2024). Esse desfecho já se desenhava no horizonte, pois com a predominância dos processos eletrônicos o mercado, sempre ávido por oportunidades, identificou com seus olhos de lince uma nova presa/oportunidade para proliferação de soluções tecnológicas - nem sempre guiadas pelo interesse público, mas pela lógica implacável da inovação orientada ao lucro. Então, emergem as *lawtechs* (empresas dedicadas ao desenvolvimento de soluções digitais para otimizar a prática jurídica).

As ferramentas visuais se consolidaram como um recurso estratégico de excelência para aprimorar a comunicação jurídica, facilitando a compreensão dos documentos e tornando-os mais acessíveis. Nesse cenário, o *visual law* desponta como uma abordagem inovadora, acompanhando o crescimento das *lawtechs* ao aliar *design*, tecnologia e linguagem simplificada. Essa convergência busca tornar o direito mais intuitivo e eficiente em um ambiente digital cada vez mais dinâmico.

O *legal design* pode ser traduzido como design jurídico, pois envolve a integração do direito com as técnicas de *design*, buscando aprimorar a

comunicação e a experiência dos usuários no contexto jurídico (Nunes; Rodrigues, 2020). Sua origem se reporta ao *design thinking*, uma abordagem concebida para o desenvolvimento de projetos voltados à resolução de problemas complexos e burocráticos, tendo a criatividade como eixo central do processo inovador e facilitador. O conceito ganhou destaque em 2009 com o lançamento do livro “*Design Thinking: uma metodologia poderosa para decretar o fim das velhas ideias*”, de Tim Brown, fundador da consultoria de inovação IDEO, no Vale do Silício, já reconhecida por sua abordagem diferenciada (Nunes; Rodrigues, 2020).

Dentre as técnicas que compõem o *legal design*, destaca-se o *visual law* (Direito Hipermodal) por sua ampla aplicabilidade em documentos jurídicos, oferecendo maior clareza e acessibilidade (Nunes; Rodrigues, 2020). Ao mesclar elementos visuais com a tradição escrita, a comunicação jurídica passa a utilizar uma gama diversificada de recursos, como infográficos, linhas do tempo, vídeos e outros elementos interativos, suavizando a necessidade da leitura de textos longos através de elementos “mais pedagógicos” que visam compreensão dos fatos em menos tempo. Trata-se da integração de técnicas que convergem a linguagem escrita com a visual e a audiovisual, possibilitada pelos avanços tecnológicos e pelos novos meios disponíveis aos operadores do Direito (Nunes; Rodrigues, 2020). É uma leitura hipermodal intrínseca ao novo modo social, ressignificando a experiência interpretativa. Considerando que devido à sobrecarga do Judiciário, textos extensos geram desconforto e comprometem a plena exegese do caso (Nunes; Paonelli, 2021).

Atualmente, diversas plataformas de fácil acesso permitem a utilização dessa abordagem visual, entre as quais se destacam o Word, com seu sistema de SmartArt, o PowerPoint, o Canvas, o Prezi, o Animaker e o Gifup (para criação de gifs) (Nunes; Rodrigues, 2020). Contudo, para preservar as terminologias já consolidadas no âmbito da linguagem e da semiótica, propõe-se que seja compreendido como direito hipermodal ou hipermodalidade jurídica (Nunes; Rodrigues, 2020). Por efeito de reconhecer que os termos tradicionais carregam um significado preciso e historicamente fundamental, ao mesmo tempo, a ideia de hipermodalidade jurídica amplia esse horizonte já consagrado.

No Brasil, a utilização da animação como meio de prova tem se tornado cada vez mais recorrente, destacando-se, entre os diversos casos, o de Isabella Nardoni, no qual essa tecnologia foi empregada por uma empresa especializada (Nunes; Rodrigues, 2020). Nesse contexto, a animação

forense serve como uma ferramenta essencial para a reconstrução dos fatos, oferecendo uma representação visual clara de eventos, o que facilita a compreensão do júri e do público em geral. No julgamento relacionado ao trágico incêndio da Boate Kiss, localizada na cidade de Santa Maria, no Rio Grande do Sul, essa técnica também foi fundamental para tentar ilustrar o caos vivido pelas vítimas nas tentativas enfrentadas para escapar, ajudando na exposição das falhas de segurança e nas responsabilidades dos acusados.

Vivencia-se uma autêntica revolução tecnológica no âmbito jurídico e processual, que não só redefine os instrumentos utilizados, mas também transforma a racionalidade do sistema de justiça, seus institutos e a atuação de seus agentes e participantes (Nunes; Paonelli, 2021). Dada a substituição da pluridimensionalidade física por petições minimalistas (mais coloridas e projetadas para simplificar o juridiquês, tornando-o visualmente mais acessível), bem como pela implementação de despachos online e julgamentos virtuais.

Ao unir pontes e pontos, a Associação Brasileira de *Lawtechs* e *Legaltechs* (AB2L) tem diversas *startups* de *Online dispute resolution* (ODR) que estão afiliadas. As ODRs emergiram tão rapidamente quanto nossa adaptação ao meio digital e, hoje, representam uma alternativa ágil e eficaz ao modelo tradicional de jurisdição estatal (Ribeiro, 2023). Seja por meio de inteligência artificial, automação ou plataformas digitais especializadas, essas ferramentas visam a eficiência dos procedimentos, reduzindo custos e prazos. Dessa forma, fala-se que essa tecnologia representa a “quarta parte” integrante dos meios consensuais de resolução de conflitos. Em razão de que a “terceira parte” é composta pelo mediador, conciliador, árbitro ou assessor, cuja função é auxiliar as partes envolvidas no conflito (Barros, 2022). E a primeira parte é representada pelo autor da demanda, enquanto a segunda é constituída pelo réu.

A Resolução n° 354, de 19 de novembro de 2020, do CNJ merece especial destaque, pois ao regulamentar as audiências virtuais de mediação e conciliação, consolida o uso das ODRs no âmbito institucional (Ribeiro, 2023). Tal medida pode ser o ponto de partida para (re)pensar o formalismo excessivo presente na arquitetura jurisdicional, promovendo um movimento mais profundo de reconfiguração das práticas jurídicas, sem desconsiderar os princípios fundamentais do direito. Posteriormente, impulsionado pela iniciativa de grandes players multinacionais, como *eBay*, *Amazon* e *Airbnb*, que implementaram mecanismos de ODR em

suas próprias plataformas de negociação, o Conselho Nacional de Justiça editou, em 2020, a Resolução nº 358 que estabeleceu o prazo de 18 meses para que os tribunais nacionais disponibilizassem um sistema informatizado para a resolução de conflitos por meio da conciliação e mediação (SIREC) (FGV, 2023).

O Conselho, além de determinar que os Tribunais desenvolvam sistemas de ODR para integrar os Centros Judiciários de Solução de Conflitos (CEJUSC), com o objetivo de utilizar a arquitetura online para promover a mediação e a conciliação, estabelece diretrizes como: a possibilidade de auditoria no código fonte e nas bases de dados, bem como a exigência de cumprimento dos princípios de segurança da informação e proteção de dados (Lei n.º 13.709/2018) (FGV, 2023). Melhor dizendo, o código-fonte é composto por um conjunto de instruções, palavras e símbolos criados por programadores, os quais definem o funcionamento de um sistema que, por seu lado, dispõe de uma base de dados, a qual engloba o conjunto de informações armazenadas e estruturadas. Desse modo, o monitoramento se torna essencial para garantir a integridade e segurança do programa.

Na cadência de que há uma complementaridade entre a solução adjudicada, característica do Poder Judiciário e as soluções não adjudicadas (Luchiari, 2012). A coexistência desses mecanismos sugere que essas duas formas de resolução de conflitos não são mutuamente exclusivas, mas sim interdependentes e podem coexistir de maneira harmônica. Nesse contexto, almeja-se um distanciamento preciso dessa visão liberal clássica, na qual a lei é vista como o único mecanismo de proteção no confronto entre a sociedade civil e o poder estatal, tendo a liberdade política do povo como o principal alvo dessa proteção (Silveira, 2020).

O acesso à justiça deve ser entendido como a obtenção de uma solução jurídica equitativa, e não apenas como o direito de acessar o Poder Judiciário ou obter uma sentença (FGV, 2023). A ênfase recai sobre a efetiva resolução do problema, seja por meios judiciais ou alternativos. Assim, a jurisdição deixa de ser a principal via de resolução de conflitos, passando a ser considerada a última *ratio* (FGV, 2023). Nos últimos anos, diversas plataformas de Resolução de Disputas Online (ODRs), que variam em modelos e funções, têm se destacado no cenário jurídico. A seguir, apresenta-se algumas localizadas por meio de uma pesquisa realizada em tribunais, evidenciando a utilização crescente de tecnologias na resolução de disputas. Os programas foram divididos em dois grupos: os modelos

gerais, mais versáteis, que oferecem soluções para diversos tipos de litígios; e os modelos específicos, que se concentram em nichos particulares, como questões de consumo e familiares.

Quadro 03: Algumas plataformas de modelo geral em operação no Judiciário nacional.

PLATAFORMA MODELO GERAL	CRIAÇÃO FUNÇÃO
SEMPROCESSO	Instituída em 2016 pelo advogado Bruno Feigelson e quatro sócios, a plataforma fornece um espaço dinâmico e inovador, permitindo que advogados de partes distintas conduzam negociações e formalizem acordos jurídicos, mitigando a necessidade de intervenção judicial.
PROJURIS (ANTIGO JUSTTO)	Concebida pelo Grupo Softplan, uma das mais proeminentes empresas de software da América Latina, a plataforma aprimora a resolução de disputas judiciais e extrajudiciais para empresas.
MEDIAR 360	Criada em 2012 por Andrea Maia (advogada, professora e palestrante de renome internacional), a plataforma constitui um ecossistema vanguardista dedicado à negociação, conciliação e mediação, oferecendo soluções céleres e eficazes para uma ampla gama de conflitos.

Fonte: Autora (2024)

Quadro 04: Algumas plataformas de modelo específico em operação no Judiciário nacional.

PLATAFORMA MODELO ESPECÍF	CRIAÇÃO FUNÇÃO
CONSUMIDOR.GOV	Instituída pelo governo brasileiro em 13 de setembro de 2014 e concebida pela Secretaria Nacional do Consumidor (Senacon), a plataforma oferece aos consumidores um canal direto e estruturado para a resolução de conflitos com empresas, promovendo o aprimoramento das relações de consumo, atenuando a judicialização e fomentando uma comunicação mais eficaz entre as partes.

PLATAFORMA MODELO ESPECÍF	CRIAÇÃO FUNÇÃO
MOL (MEDIAÇÃO ONLINE)	Fundada por Melissa Felipe Gava (advogada e CEO) e Camilla Feliciano Lopes , a plataforma foi laureada com o prêmio “Conciliar é Legal” , concedido pelo Conselho Nacional de Justiça. Tendo em vista que apresenta uma solução tecnológica desenvolvida para a mediação de reclamações em órgãos como Procon, Reclame Aqui, Consumidor.gov e Banco Central do Brasil.
RESOLVVI	Idealizada em 2017, trata-se de uma empresa de tecnologia especializada em auxiliar consumidores brasileiros na reivindicação de indenizações por transtornos com voos e negativas indevidas. Por meio de uma plataforma digital segura e desprovida de burocracia, viabiliza um processo ágil, eficiente e acessível.
OS NOSSOS	Em 2020, inspirado no conceito das famílias mosaico (“os meus, os seus e os nossos” (filhos)), surgiu o aplicativo que foi desenvolvido para auxiliar famílias nos desafios pós-divórcio. Dessa forma, oferece soluções para a organização da vida familiar e melhora a comunicação entre os pais, promovendo um ambiente de maior segurança e estabilidade para os filhos.
ADAM TECNOLOGIA	Desde 2001 a empresa se dedica ao desenvolvimento de soluções inovadoras para instituições e profissionais que conduzem processos de Conciliação, Mediação e Arbitragem. Diferencia-se ao oferecer aos usuários total autonomia para criar plataformas personalizadas de acordo com suas necessidades, permitindo a configuração de páginas, imagens e etc.

Fonte: Autora (2024)

Ao permitir que as partes envolvidas resolvam questões sem recorrer diretamente ao sistema judiciário tradicional, essas ferramentas ajudam a diminuir a sobrecarga dos tribunais e oferecem uma alternativa prática, especialmente em situações em que o litígio formal seria desproporcional ao valor da disputa. Esses softwares representam um movimento significativo em direção à modernização e desburocratização da resolução de conflitos, alavancando o poder da tecnologia para transformar as disputas que são tratadas no Brasil.

Como um claro reconhecimento o Conselho Nacional de Justiça concede o prêmio “Conciliar é Legal” para iniciativas inovadoras como a MOL (Mediação Online), que se destacam pela criatividade e relevância na promoção de soluções alternativas para a mediação e conciliação. O Prêmio Conciliar é Legal reconhece, valoriza, divulga e incentiva iniciativas de modernização no Poder Judiciário que favoreçam a aproximação entre as partes, promovam a pacificação eficaz por meio da resolução de conflitos e, assim, contribuam para o aprimoramento da Justiça (CNJ, 2024).

Em 2023, 45 dos 91 tribunais brasileiros aderiram a essa importante iniciativa, conforme os dados divulgados no site do Conselho Nacional de Justiça (CNJ). De acordo com as informações disponíveis no painel, um quadro comparativo foi criado, detalhando a participação dos tribunais de diferentes segmentos da Justiça.

Quadro 05: Relação de tribunais brasileiros que implementaram plataformas digitais de resolução de conflitos.

TRIBUNAL	REGIÃO
Tribunais Estaduais (21)	TJAP – Tribunal de Justiça do Amapá TJCE – Tribunal de Justiça do Ceará TJES – Tribunal de Justiça do Espírito Santo TJGO – Tribunal de Justiça de Goiás TJMA – Tribunal de Justiça do Maranhão TJMG – Tribunal de Justiça de Minas Gerais TJMS – Tribunal de Justiça de Mato Grosso do Sul TJMT – Tribunal de Justiça de Mato Grosso TJPB – Tribunal de Justiça da Paraíba TJPE – Tribunal de Justiça de Pernambuco TJPI – Tribunal de Justiça do Piauí TJPR – Tribunal de Justiça do Paraná TJRJ – Tribunal de Justiça do Rio de Janeiro TJRN – Tribunal de Justiça do Rio Grande do Norte TJRO – Tribunal de Justiça de Rondônia TJRR – Tribunal de Justiça de Roraima TJRS – Tribunal de Justiça do Rio Grande do Sul TJSC – Tribunal de Justiça de Santa Catarina TJSE – Tribunal de Justiça de Sergipe TJSP – Tribunal de Justiça de São Paulo TJTO – Tribunal de Justiça do Tocantins
Tribunais Regionais Federais (05)	TRF1 – Tribunal Regional Federal da 1ª Região TRF2 – Tribunal Regional Federal da 2ª Região TRF3 – Tribunal Regional Federal da 3ª Região TRF4 – Tribunal Regional Federal da 4ª Região TRF5 – Tribunal Regional Federal da 5ª Região

TRIBUNAL	REGIÃO
Tribunais Regionais do Trabalho (19)	TRT1 – Tribunal Regional do Trabalho da 1ª Região
	TRT2 – Tribunal Regional do Trabalho da 2ª Região
	TRT3 – Tribunal Regional do Trabalho da 3ª Região
	TRT4 – Tribunal Regional do Trabalho da 4ª Região
	TRT5 – Tribunal Regional do Trabalho da 5ª Região
	TRT6 – Tribunal Regional do Trabalho da 6ª Região
	TRT7 – Tribunal Regional do Trabalho da 7ª Região
	TRT10 – Tribunal Regional do Trabalho da 10ª Região
	TRT12 – Tribunal Regional do Trabalho da 12ª Região
	TRT13 – Tribunal Regional do Trabalho da 13ª Região
	TRT15 – Tribunal Regional do Trabalho da 15ª Região
	TRT16 – Tribunal Regional do Trabalho da 16ª Região
	TRT19 – Tribunal Regional do Trabalho da 19ª Região
	TRT20 – Tribunal Regional do Trabalho da 20ª Região
	TRT21 – Tribunal Regional do Trabalho da 21ª Região
	TRT22 – Tribunal Regional do Trabalho da 22ª Região
	TRT23 – Tribunal Regional do Trabalho da 23ª Região
	TRT24 – Tribunal Regional do Trabalho da 24ª Região

Fonte: Autora (2024)

A ampla adesão de tribunais de todo o país, abrangendo diferentes segmentos da Justiça, reflete uma mudança cultural no Poder Judiciário, que busca soluções mais rápidas e eficazes para a resolução de disputas, aliviando a sobrecarga do sistema e evitando a judicialização excessiva. A aplicação da inteligência artificial na resolução de conflitos online viabiliza a criação de sistemas computacionais preditivos que, por meio do processamento de grandes volumes de dados, são aptos a fornecer ao operador uma previsão sobre a probabilidade de uma disputa ser solucionada por conciliação ou mediação (Faria, s.d).

Ademais, detém a aptidão de identificar a técnica de autocomposição mais apropriada à resolução do caso concreto, além de antever o tempo de negociação essencial para que as partes alcancem um consenso (Faria, s.d). A esse respeito, adverte-se que a escolha do procedimento mais eficaz depende de diversos fatores, como a natureza do conflito, o grau de litigiosidade entre as partes e a complexidade das demandas envolvidas.

Avançando ainda mais, há, em tese, a promessa de que o sistema preditivo integre outras tecnologias e distintas metodologias, como a Teoria dos Jogos, para o desenvolvimento de sistemas capazes de operar com dados extraídos do subconsciente das partes, identificando padrões

de comportamento humano e ampliando as chances de se alcançar um consenso na resolução do litígio (Faria, s.d). A ideia central é a de que, ao incorporar elementos de observação comportamental e preditiva, esses sistemas poderiam antecipar reações e tendências que, muitas vezes, não são explicitamente manifestadas durante a negociação, que influenciam as negociações de maneira sutil.

Nesse patamar mais elevado de desenvolvimento, os robôs dotados de inteligência artificial serão aptos a mediar conflitos de forma autônoma, discernindo o ambiente em que operam por meio da identificação de padrões comportamentais das partes, das particularidades do problema, do direito aplicável, entre outros parâmetros (Faria, s.d). A redação dá a entender a capacidade dos robôs para interpretar o ambiente em que operam como um alquimista digital, capaz de transformar dados brutos em *insights* precisos e, assim, desvendar os mistérios ocultos das interações humanas.

Por último, mas não menos importante, o Judiciário, longe de ser uma panaceia, trouxe à baila pelo Código de Processo Civil de 2015 o Incidente de Resolução de Demandas Repetitivas (IRDR). Inspirado no direito estrangeiro, especialmente no modelo alemão do *Musterverfahren* (procedimento-padrão), o mecanismo busca incentivar a uniformização da jurisprudência e enfrentar a escalada exponencial de processos. Já que a procura por prestação jurisdicional é extraordinária e excede a capacidade de fornecer uma resposta qualitativa e tempestivamente satisfatória (Mendes, 2017).

Impõe-se a racionalização do processo judicial, para resolver múltiplas questões semelhantes de forma única, em vez de perpetuar julgamentos fragmentados sobre temas idênticos. O mecanismo diante das demandas repetitivas deverá ceder à lógica da gestão, na qual o julgamento incessante e descontrolado dará lugar à lógica da racionalização, em que, como se afirma hoje, o “menos pode ser mais” (Mendes, 2017).

O Tribunal Regional Federal da 5ª Região divulgou um levantamento minucioso sobre a quantidade de IRDRs processados e julgados em sua competência. O documento em questão apresenta o total de IRDRs suscitados no TRF5 entre 2016 e 01 de outubro de 2024, detalhando a quantidade de incidentes registrados a cada ano nesse intervalo (TRF5, 2024). O relatório contempla dados acerca da admissibilidade, dos temas abordados e das decisões proferidas, oferecendo uma visão completa sobre a aplicação do IRDR na jurisdição do tribunal e o número total de IRDRs

suscitados alcançou 24 casos até 2024, com variações significativas na quantidade de incidentes em cada ano.

Por sua parte, o Tribunal de Justiça do Estado do Pará também disponibilizou um relatório quantitativo e qualitativo apresentando os incidentes gerados na região. O demonstrativo considera todos os incidentes suscitados (pendentes de admissibilidade, inadmitidos, admitidos, julgados com a interposição de recurso extraordinário e/ou especial e com trânsito em julgado) desde a vigência do novo Código de Processo Civil, em 2015, até o mês de maio de 2021 (TJPA, 2021). E a quantidade total de IRDRs alcançou 18 casos até o ano de 2021.

Assim, independentemente de sua origem, é inegável que o instituto do IRDR (Incidente de Resolução de Demandas Repetitivas) constitui uma realidade na legislação processual brasileira (Mendes, 2017). Seu principal propósito é proporcionar maior celeridade e assegurar a isonomia no julgamento de causas repetitivas, o que se revela benéfico para o magistrado contemporâneo, a quem se exige não apenas competências técnico-jurídicas, mas também habilidades de gestão e administração (Mendes, 2017). No entanto, é imprescindível cautela para que não se perca a essência do direito, que deve ser a justiça, em nome de um modelo que poderia transformar o judiciário em uma “linha de produção” de decisões, muitas vezes sem a devida atenção à complexidade de cada caso.

Em face do exposto, avista-se uma nova forma de “realidade processual” de matriz (tecno)liberal que, sob o pretexto da eficiência, pode obscurecer a substância hermenêutica do direito. Essa tendência expansiva da tecnificação jurídica não se limita à estrutura judiciária, mas alcança também sua dimensão epistemológica e formativa. A gamificação do ensino jurídico, ao recorrer a estratégias lúdicas para simplificar a compreensão do funcionamento do Judiciário, traduz uma tentativa de adaptação pedagógica ao espírito dos nativos digitais (Y, Z e alpha), os quais já possuem a lógica da instantaneidade de forma intrínseca. Porquanto, as interações humanas deixaram de ocorrer por meio dos átomos e passaram a se desenvolver a partir dos bits (Faria, 2022).

Do ponto de vista operacional, a gamificação teve como inspiração o jogo Mario Party, da Nintendo, configurando-se como uma ampla gincana competitiva composta por diversos minigames, cuja dinâmica resulta na atribuição de pontuações (Bolesina; Gervasoni, 2021). Não obstante, o direito, em sua essência, exige pensamento crítico e interpretação, competências que podem ser minimizadas quando se priorizam aspectos

competitivos e de recompensa imediata. A automatização do direito corre o risco de transformar a justiça em um simulacro, onde a previsibilidade estatística substitui a complexidade interpretativa e a busca pelo justo.

O ser humano acaba por se inferiorizar diante da potência da tecnologia e acaba (des)construindo o seu destino. Nós, seres humanos, estamos progressivamente perdendo nossa consciência, enquanto as máquinas começam a despertar (Weeb, 2020). A mentalidade “*nowist*” valoriza as conquistas tecnológicas imediatas, mas exime da responsabilidade de refletir sobre a evolução da tecnologia e as repercussões e consequências de nossas futuras ações (Weeb, 2020). O Oráculo, personagem do filme Matrix, reconheceu que o comportamento humano é guiado pelas escolhas, e não pela simplificação da relação binária de causa e efeito.

Então, antes que seja tarde, uma contemplação profunda, firme e atenta sobre os benefícios e riscos da inteligência artificial se faz essencial, pois ao negligenciar as implicações a longo prazo das escolhas do período presente, abre-se a margem para a humanidade ser subjugada pelas próprias criações, invertendo a ordem das coisas. Tal como o monstro de Frankenstein, em sua eloquente manifestação, diria ao seu criador: “Você me deu a vida, mas agora sou eu quem detém o poder sobre você.”

4 INTELIGÊNCIA ARTIFICIAL E DECISÃO JUDICIAL: RISCOS E BENEFÍCIOS NO USO DA IA

“Você sente que tem algo errado com o mundo, mas não sabe o que é. Está lá, como uma farpa em sua mente”

– Morpheus, em Matrix (Wachowski; Wachowski, 1999)

No primeiro capítulo, ao percorrermos a trajetória histórica da inteligência artificial, foi possível observar como seu desenvolvimento se entrelaça com grandes debates filosóficos, psicológicos e linguísticos. Desde as primeiras reflexões sobre a relação entre razão e emoção até a consolidação da IA como um campo autônomo da ciência da computação, tornou-se evidente que a inteligência artificial não apenas reflete, mas também ressignifica questões fundamentais sobre a própria natureza do pensamento e da tomada de decisão.

No contexto do direito, essas questões assumem especial relevância. A decisão judicial não é um simples exercício lógico, mas um processo interpretativo que envolve nuances linguísticas, valores e subjetividades. A IA, ao ser integrada a esse universo, promete benefícios notáveis, como maior eficiência e padronização. No entanto, também suscita preocupações sobre o que esperar da justiça do futuro, sobretudo quanto ao risco de decisões automatizadas, além das inquietações sobre a possível substituição de profissionais da área jurídica. Se já se fala na contingência de um “juiz robô”, o que restará para as demais funções essenciais ao direito?

Umberto Eco, dividiu a sociedade entre os apocalípticos, que veem a cultura de massa como uma ameaça à alta cultura e ao pensamento crítico, e os integrados, que a enxergam como uma forma de democratização do conhecimento e do entretenimento. Ainda, menciona que o apocalipse é a obsessão do dissidente, enquanto a integração é a realidade concreta daqueles que não discordam (Eco, 1984). Essa dicotomia entre temor e aceitação reflete o próprio debate sobre a inserção da inteligência artificial no direito, onde os apocalípticos alertam para os perigos da automatização excessiva, que pode perpetuar vieses, comprometer a transparência das decisões judiciais e reduzir a sensibilidade humana na aplicação da justiça, enquanto, os integrados enxergam a IA como uma ferramenta de

aprimoramento, capaz de tornar o sistema jurídico mais ágil, acessível e eficiente, sem necessariamente substituir o papel do jurista.

Assim, este capítulo buscará levantar o véu dos mistérios sobre os riscos e benefícios da aplicação da inteligência artificial no processo judicial, refletindo sobre os impactos dessa tecnologia na busca por um equilíbrio entre precisão algorítmica e sensibilidade humana, visto que um julgamento conduzido com irreflexão e precipitação não apenas compromete a justiça, mas também fragiliza a confiança no próprio sistema jurídico.

Aristóteles declarou que a “alma racional”, a capacidade de raciocinar, fazer abstrações e refletir, era a singularidade humana (Bennett, 2024). O ser humano é uma espécie dotada de inteligência, pensamentos e senso moral especial, o que torna a interação entre a tecnologia e os aspectos subjetivos do comportamento humano ainda mais complexa. Esse aspecto da experiência humana, conhecido como *qualia*, refere-se àquelas sensações e percepções subjetivas únicas, como sentimentos. A *qualia* é um termo de origem latina (“qual tipo”) que foi introduzido na filosofia contemporânea para se referir às qualidades fenomenais, subjetivas e conscientes da nossa experiência mental, que seriam acessíveis por meio da introspecção (Pereira, 2013). Esses elementos da experiência não podem ser completamente explicados ou descritos por meio de dados objetivos ou físicos, como atividades cerebrais ou processos neurobiológicos.

Isso significa que, por mais que a ciência avance no estudo do cérebro e da mente, a experiência subjetiva continua sendo algo que escapa a explicações puramente físicas ou comportamentais. Essa característica se torna ainda mais acentuada quando tentamos aplicá-la na inteligência artificial: uma máquina poderia ter *qualia*, ou ela apenas simula uma experiência consciente sem realmente vivê-la? Daí, cria-se a lacuna entre a experiência subjetiva humana e a objetividade das máquinas, pois há diferença entre “simular” uma experiência e “vivê-la” de fato.

Há algo único no que significa sentir dor; há algo particular no que significa perceber a cor vermelha; há algo distinto no que significa estar apaixonado; há algo específico no que significa estar deprimido, e assim por diante (Pereira, 2013). Então, mesmo que uma máquina possa representar a percepção do vermelho ou da dor, isso não significa que ela esteja realmente experimentando essas sensações de maneira subjetiva.

Embora a inteligência artificial possa ser uma ferramenta valiosa no processamento de informações, ela não é capaz de substituir o discernimento moral e ético que é intrínseco à consciência humana e à interpretação das

leis. A IA não substituirá os operadores do direito, mas sim os “juristas 5.0” (aqueles que dominarem os fundamentos técnicos da inteligência artificial e aplicá-la de maneira adequada), superarão os profissionais que se recusam a adotá-la. A verdadeira questão, portanto, reside em como utilizá-la. Além do que as empresas atingem seu maior potencial quando humanos e máquinas colaboram como aliados, unindo suas forças complementares em vez de se colocarem como adversários (Daugherty; Wilson, 2019).

Ao longo de três séculos de descobertas e explorações, os seres humanos interpretaram o mundo como Kant havia previsto: através da estrutura da mente humana (Kissinger; Schmidt; Huttenlocher, 2023). Ou melhor, desde as revoluções científicas até a era digital, a compreensão humana da realidade é mediada por categorias cognitivas, como tempo, espaço e causalidade, que organizam nossas experiências e conhecimentos. No entanto, ao se deparar com os limites da capacidade cognitiva, tornou-se necessário recorrer às máquinas (computadores) para expandir a reflexão e superar essas barreiras (Kissinger; Schmidt; Huttenlocher, 2023).

A IA, ao oferecer respostas rápidas e simplificadas, pode sutilmente atenuar nossa inquietação intelectual e social, enfraquecendo o ímpeto natural de indagar (Premuzic, 2024). Se, por um lado, a tecnologia brinda com acesso imediato à informação, por outro, pode insidiosamente acomodar a autonomia cognitiva de reflexões profundas em virtude da dependência de respostas prontas. Os mecanismos de busca, por sua parte, recorrem à inteligência artificial para atender às consultas, nesse caso os seres humanos delegam à tecnologia aspectos essenciais da própria cognição humana (Kissinger; Schmidt; Huttenlocher, 2023).

Atualmente, as aplicações de IA nos tribunais e na prática jurídica são majoritariamente baseadas em modelos de aprendizado de máquina, que identificam padrões em documentos por meio de métodos estatísticos, servindo de base para previsões e decisões automatizadas (CNJ, 2024). Ou seja, aquilo que aparenta ser um raciocínio pautado em grandes conhecimentos é, na verdade, apenas uma correspondência de um vasto conjunto de textos. Isso significa que, salvo eventuais mecanismos auxiliares de ontologias semânticas para classificar as informações extraídas, os modelos de IA não possuem uma representação estruturada do conhecimento jurídico nem a capacidade de realizar inferências próprias do raciocínio jurídico (CNJ, 2024). A mera identificação de recorrências documentais não equivale à aplicação do direito, pois ignora a historicidade e a complexidade decisória.

O cérebro humano possui um órgão dedicado à linguagem: os seres humanos desenvolveram duas novas regiões no neocórtex, as áreas de Broca e Wernicke, interligadas em uma sub-rede especializada na linguagem (Bennett, 2024). Com base nessa complexidade biológica, lembra-se que Geoffrey Hinton, um dos padrinhos da inteligência artificial, sempre defendeu que a única maneira de fazer a IA funcionar é aproximando sua computação do funcionamento do cérebro humano. É justamente nos dilemas sobre o mundo físico e mental que se começa a perceber as limitações dos LLMs - como eles diferem da linguagem no cérebro humano e quais aspectos da inteligência precisam ser compreendidos para a engenharia reversa, na pretensão de desenvolver sistemas de linguagem de IA que operem de maneira mais humana (Bennett, 2024).

A função da linguagem, essencial ao direito, e os aspectos psicológicos envolvidos na tomada de decisão tornam essa interseção entre IA e Judiciário um tema tão instigante quanto desafiador. Afinal, para as coisas novas, necessitam-se de palavras novas, assim o quer a clareza da linguagem para evitar a confusão inseparável do sentido múltiplo dos mesmos vocábulos (Kardec, 2021). Na Grécia Antiga, por exemplo, diferentes palavras eram usadas para descrever o amor, como “éros” para o desejo romântico, “*philia*” para a amizade e “ágape” para um amor altruísta e incondicional. Essa distinção linguística evitava ambiguidades e reforçava a necessidade de precisão conceitual.

A língua portuguesa, por sua própria natureza, ostenta uma complexidade inerente, em virtude da diversidade de significados que uma palavra ou frase pode abarcar (Sale, 2020). A polissemia do termo “pena” é um exemplo marcante, pois pode designar tanto a estrutura das aves quanto, no âmbito jurídico, a sanção imposta a quem comete um crime ou infração. Quando volta-se nossa atenção para o contexto jurídico, nota-se que o vocabulário se torna ainda mais intrincado, o que intensifica uma dificuldade linguística maior, podendo resultar em distorções e interpretações equivocadas, devido à ambiguidade da linguagem utilizada (Sale, 2023). Essa diversidade de significados na linguagem humana também respinga nos vieses dos programadores e de seus textos utilizados em seu treinamento.

É inegável que a sociedade digital avança em ritmo acelerado, exigindo que o Direito acompanhe essa transformação, aperfeiçoe-se, renove seus institutos e desenvolva novos mecanismos para assegurar a segurança jurídica nas relações sociais, do contrário, corre o risco de tornar-

se obsoleto, fomentando a prática da justiça com o próprio mouse e todas as consequências nefastas decorrentes do uso arbitrário das próprias razões, bem como do desequilíbrio instaurado pelo poder desmedido das grandes corporações, detentoras dos recursos essenciais à vida digital (Pinheiro, 2021). Além das questões já mencionadas, é essencial considerar o risco da autotutela decorrente da desatualização ou insuficiência das normas, que abre espaços para a imposição de soluções privadas sem a devida regulação. Nessa lacuna normativa, plataformas de resolução de disputas, decisões unilaterais e algoritmos sob o controle de corporações passam a definir parâmetros de justiça e mediação de conflitos, muitas vezes sem transparência ou garantias de equidade.

Então, à medida que os sistemas de IA se tornam mais avançados e autônomos, surgem questões legais complexas sobre responsabilidade e tomada de decisões. Dado que os avanços recentes em neurociência e ciência da computação favorecem o surgimento de novas abordagens do realismo jurídico (Streck; Jung, 2024). Podendo ser denominadas como realismo jurídico 2.0 ou realismo jurídico de alta tecnologia (Streck; Jung, 2024). O que se apresenta como superação da subjetividade é, na verdade, a substituição de uma racionalidade interpretativa por uma racionalidade estatística (igualmente cega), porém travestida de neutralidade científica.

Destarte, sugere-se que, no âmbito do desenvolvimento de sistemas de inteligência artificial, é igualmente relevante considerar não apenas o processo de criação desses sistemas, mas também o processo de aplicação ética diante do elemento humano que os utiliza (Pinto, 2020). Ou seja, a ética não deve ser uma consideração secundária, mas sim uma parte integrante de todo o processo, desde a concepção até o uso prático. Em razão de que o armazenamento e o processamento de dados pessoais devem estar em conformidade com as diretrizes estabelecidas pelas regulamentações presentes no ordenamento jurídico.

Ao passo que, do ponto de vista da Lei Geral de Proteção de Dados (LGPD), é factível a alegação de que os sistemas de inteligência artificial devem ser programados de modo a aderir aos dez princípios delineados no artigo 6º desta legislação, dado que essas ferramentas estão envolvidas no processamento de vastas quantidades de informações (Sale, 2023). Isso indica que a LGPD estabelece diretrizes rigorosas para a coleta, processamento e uso de dados pessoais, e a inteligência artificial lida com uma quantidade massiva de informações, muitas vezes contendo dados sensíveis.

Entre os princípios listados, merecem destaque, quando se trata da avaliação dos processos automatizados, aqueles referentes à finalidade, transparência, segurança, prevenção e não discriminação, visto que representam diretrizes desafiadoras para os profissionais que operam e programam as tecnologias relacionadas à inteligência artificial (Sale, 2023). Devido que as IAs “aprendem” ao processar dados e posteriormente elaboram suas observações e conclusões com base neles (Kissinger; Schmidt; Huttenlocher, 2023). Com essa conjuntura, vê-se que elas carecem da capacidade de compreensão profunda porque seu funcionamento se limita a entender informações, mas não a verdadeiramente compreendê-las de maneira contextual e abrangente.

Assim, a esta atmosfera de incertezas, acresce uma inquietação quanto à salvaguarda dos dados pessoais, que são componentes cruciais para o adequado funcionamento dos sistemas de tomada de decisões automatizadas, pois é por meio dos dados que as ferramentas de inteligência artificial conduzem suas verificações para efetuarem os processos decisórios (Melo, 2023). Na busca do equilíbrio entre inovação e prudência, o Conselho Nacional de Justiça estabeleceu diretrizes para o uso da IA no sistema judiciário por meio da Resolução nº 332, de 21 de agosto de 2020, estabeleceu as primeiras diretrizes sobre a ética, a transparência e a governança na criação e aplicação da Inteligência Artificial no Poder Judiciário, além de outras disposições.

Em seu artigo 5º é lecionado que a aplicação de modelos de inteligência artificial deve assegurar a segurança jurídica e contribuir para que o Poder Judiciário mantenha a igualdade de tratamento em casos idênticos. Na sequência, tem-se que as decisões judiciais baseadas em ferramentas de IA devem garantir a igualdade, a não discriminação, a pluralidade e a solidariedade, promovendo um julgamento justo e criando condições para eliminar ou reduzir a opressão, a marginalização humana e os erros de julgamento resultantes de preconceitos.

E, antes de sua implementação, o modelo de IA deve passar por um processo de homologação para identificar possíveis preconceitos ou generalizações que possam ter influenciado seu desenvolvimento, resultando em tendências discriminatórias em seu funcionamento. Caso não seja possível eliminar o viés discriminatório do modelo de Inteligência Artificial, sua utilização deverá ser descontinuada. Para mais, os usuários externos devem ser claramente informados, em linguagem acessível

e precisa, sobre o uso de sistemas inteligentes nos serviços que lhes são oferecidos.

A Unesco publicou o documento “Diretrizes da UNESCO para o uso de sistemas de IA em tribunais” que trata sobre desafios e as oportunidades relacionados ao uso ético, transparente e responsável de sistemas de IA na área jurídica, incluindo sua aplicação na elaboração de documentos legais, na construção de argumentos e na prolação de decisões. Em decorrência que a integração de sistemas de inteligência artificial aos processos decisórios podem causar um efeito cascata em todo o sistema judicial (UNESCO, 2024).

Por um lado, a utilização adequada desses sistemas pode favorecer o desempenho das atividades desempenhadas pelos servidores judiciais e ao mesmo tempo, sob outra vertente os efeitos resultantes da aplicação da IA podem ser negativos, notadamente quando reforçam desigualdades estruturais por meio de discriminações automatizadas (UNESCO, 2024). Além de que a linguagem figurada pode gerar equívocos de interpretação. As ferramentas de inteligência artificial não possuem a capacidade de substituir o raciocínio jurídico qualificado, tampouco podem suplantar o discernimento humano ou a orientação jurídica personalizada (UNESCO, 2024).

No que se refere às IAGs, as diretrizes da Unesco apresentam: (i) recomendações de governança voltadas às instituições do Judiciário e (ii) orientações específicas destinadas a membros individuais que integram o sistema judicial (CNJ, 2024). Na primeira parte, destaca-se que para garantir o uso ético e responsável da inteligência artificial generativa no âmbito do Poder Judiciário, é fundamental investir na conscientização de magistrados e servidores sobre as funcionalidades, limitações e riscos dessas tecnologias.

Além disso, torna-se essencial assegurar a autenticidade e a integridade dos conteúdos produzidos por IA, mediante rotulagem clara, rastreamento de alterações e a implementação de protocolos de certificação que atendam às exigências legais e às diretrizes nacionais e internacionais. E é imprescindível estabelecer restrições claras quanto ao uso dessas ferramentas, especialmente em contextos que possam comprometer direitos humanos ou afetar a integridade do processo judicial, assegurando que sua aplicação ocorra dentro de limites éticos e jurídicos bem definidos.

Na parte específica voltada a membros individuais do Judiciário, destaca-se a importância de uma postura cautelosa frente ao uso de

ferramentas de IAGs, sobretudo porque os termos de uso adotados pela maioria das empresas que oferecem acesso gratuito a estas tecnologias frequentemente autorizam a utilização das informações inseridas pelos usuários para o treinamento de novos modelos, por isso, deve-se evitar inserir dados pessoais ou informações confidenciais em plataformas públicas. Orienta-se, ainda, que as saídas (respostas) produzidas por ferramentas de IA sejam confrontadas com fontes confiáveis tendo em vista a chance de imprecisões, erros factuais ou distorções.

Sob o prisma operacional, os modelos de linguagem de uso geral, disponibilizados comercialmente, não devem ser considerados fontes fidedignas de informação, tampouco instrumentos apropriados para a condução de investigações jurídicas ou a execução de tarefas matemáticas. Ainda que se utilizem comandos cuidadosamente elaborados, as respostas produzidas podem ser imprecisas, incompletas ou enganosas. Com isso, é imprescindível que os operadores do direito estejam plenamente conscientes de que as respostas produzidas por modelos de linguagem de larga escala podem conter informações incorretas, imprecisas ou até mesmo fictícias, abrangendo aspectos factuais, jurídicos e técnicos, além de envolver riscos éticos como a reprodução de vieses e a violação de direitos relacionados à privacidade.

A construção articulada e a aparência de coerência dos textos gerados por modelos de linguagem não devem ensejar uma confiança absoluta quanto à veracidade ou à adequação do conteúdo apresentado. É, pois, imperativo que todo conteúdo oriundo desses sistemas seja submetido a uma revisão rigorosa e criteriosa antes de ser incorporado a documentos jurídicos formais, como peças processuais, despachos, votos ou sentenças. Em consonância com esse cuidado, utilização de sistemas de inteligência artificial generativa deve ser devidamente sinalizada quando utilizada na elaboração de textos jurídicos ou audiências judiciais, de modo a preservar direitos das partes. Para mais, no que tange a integridade o documento enfatiza que as produções textuais geradas pelos sistemas de IAG não devem ser indevidamente equiparados a produções humanas autênticas, sob pena de possível violação de direitos autorais ou de propriedade intelectual decorrente de sua utilização.

A Resolução nº 615, de 11 de Março de 2025 refinou e expandiu as diretrizes, regulamentando de forma mais detalhada o desenvolvimento e governança das ferramentas de IA. Posto que a capacidade de adaptação do Direito é fundamental para a própria segurança do ordenamento,

assegurando a estabilidade do sistema por meio da atuação legítima do poder encarregado de produzir normas válidas e eficazes (Pinheiro, 2021).

De modo particular, o Capítulo VI lançou luz ao disciplinar o uso e a contratação de Modelos de Linguagem de Larga Escala (LLMs) e outros Sistemas de IA Generativa (IAGen), ampliando a compreensão e a governança sobre essas tecnologias no Judiciário. Na redação do artigo 19, determina-se que os usuários deverão participar de capacitações e treinamentos específicos sobre melhores práticas, limitações, riscos e o uso ético, responsável e eficiente de LLMs e sistemas de IA generativa em suas atividades, seguindo um programa padronizado de letramento digital.

Em seguida, asseverou-se que o uso dessas ferramentas será de caráter auxiliar e complementar, consistindo em mecanismos de apoio à decisão, vedada a utilização como instrumento autônomo de tomada de decisões judiciais sem a devida orientação, interpretação, verificação e revisão por parte do magistrado, que permanecerá integralmente responsável pelas decisões tomadas e pelas informações nelas contidas. E, inclusive, foi atestado que os sistemas contratados deverão incorporar os princípios de *privacy by design* (privacidade desde a concepção) e *privacy by default* (privacidade por padrão), garantindo, sempre que necessário, a opção de não armazenar ou eliminar o histórico de perguntas e *prompts* (comandos) e poderá ser exigido um relatório com indicadores claros para avaliar sua implementação e conformidade. Desse modo, já não basta dominar apenas o Direito e a legislação; é essencial compreender os modelos que os orientam (Pinheiro, 2021).

À medida que as estruturas e sistemas que regem a sociedade se tornam cada vez mais dependentes da IA, perceberemos que as decisões tomadas em nosso nome farão perfeito sentido para as máquinas, mas nem sempre para uma pessoa (Webb, 2020). A desconexão entre a lógica algorítmica e a complexidade da experiência humana revela que o que é racional para uma máquina pode não ser necessariamente justo ou adequado para os indivíduos afetados.

Os 09 titãs da tecnologia, como *Google, Amazon, Apple, IBM, Microsoft e Facebook*, nos Estados Unidos, e *Baidu, Alibaba e Tencent*, na China, estão desenvolvendo a inteligência artificial com o objetivo de construir um futuro mais promissor e positivo para a sociedade (Webb, 2020). No entanto, essa promessa pode ter intenções veladas como possibilidade de um sistema de monitoramento global semelhante ao déjà vu em Matrix - um sinal quase imperceptível de que algo foi alterado na

estrutura do mundo, mas que passa despercebido pela maioria. Como no filme, onde a repetição de um evento indica manipulação oculta do sistema, a crescente influência dessas corporações pode modificar silenciosamente as regras do jogo, moldando a realidade sem que a sociedade perceba a mudança até que seja tarde demais. Assim, compreende-se que a IA vai além de um mero produto concebido no Vale do Silício para ser comercializado enquanto o mercado se mantém aquecido (Webb, 2020).

Embora o número de indivíduos capacitados para desenvolver inteligência artificial esteja em constante crescimento, ainda é alarmantemente reduzida a parcela daqueles que se dedicam a refletir sobre suas implicações para a humanidade, seja sob a perspectiva social, jurídica, filosófica, espiritual ou moral (Kissinger; Schmidt; Huttenlocher, 2023). Se poucos se dedicam a questionar como essa tecnologia afeta direitos fundamentais, relações sociais e até mesmo a concepção do que significa ser humano, abre-se margem para consequências imprevistas e potencialmente irreversíveis. Há entusiasmo, mas também descompasso. A utilização de algoritmos que neutralizam ou sequestram nossa atenção está contribuindo para uma crise de distração (Premuzic, 2024).

O filósofo e polímata francês Montesquieu afirmou: “Nossa alma foi concebida para pensar, ou seja, para perceber; e um ser dotado dessa capacidade deve necessariamente possuir curiosidade, pois como todas as coisas estão interligadas em uma cadeia, onde cada ideia decorre de uma anterior e antecede outra, é impossível contemplar uma sem desejar alcançar a próxima” (Kissinger; Schmidt; Huttenlocher, 2023). Em outras palavras, cada descoberta não encerra um ciclo, mas abre caminho para novas indagações.

4.1 O déjà vu a (des)serviço do direito: a implementação da IA no poder judiciário

É essencial reconhecer que a ressaca tecnológica fomenta uma relação de dependência, cujos impactos se estendem a indivíduos, empresas, governos e instituições (Pinheiro, 2021). A dependência, seja do que for, cobra seu preço. Não se trata apenas de indivíduos desorientados sem seus dispositivos, mas de empresas, governos e instituições inteiras que se veem reféns de sistemas dos quais não conseguem mais se desvencilhar. A questão, então, não é apenas o uso da tecnologia, mas o quanto nos tornamos incapazes de viver sem ela. A era da IA nos conduz a uma impaciência

crescente, a uma ignorância disfarçada de conhecimento e a uma desilusão subjacente, ao mesmo tempo em que reforça nossas interpretações egoístas do mundo (Premuzic, 2024).

Os sistemas de IA Generativa (IAG) começaram a ganhar popularidade em 2022, impulsionados pelo lançamento de ferramentas capazes de gerar imagens a partir de comandos em texto (*prompts*), porém foi entre o final de 2022 e o início de 2023 que essa tecnologia alcançou seu auge, com o lançamento do ChatGPT pela OpenAI, que rapidamente se tornou um fenômeno global (CNJ, 2024). Neste mesmo ano, em novembro de 2023, o Conselho Nacional de Justiça iniciou uma investigação sobre o juiz federal da 1ª Região que recorreu ao ChatGPT para redigir uma sentença, na qual foram incorporadas jurisprudências inexistentes do Superior Tribunal de Justiça (STJ), geradas pela ferramenta de inteligência artificial. Em outro episódio, na Colômbia, um magistrado usou a mesma ferramenta como auxílio na fundamentação de uma decisão relacionada ao direito à saúde de uma criança autista.

Os LLMs, como o GPT3, operam prevendo a próxima palavra em uma frase, sentença ou parágrafo. E ao contrário de seus predecessores, o GPT-3 foi treinado com uma quantidade massiva de dados extraídos de praticamente toda a internet (Bennett, 2024). Ao utilizar uma quantidade massiva de dados provenientes de praticamente toda a internet, o modelo adquiriu uma capacidade impressionante de gerar textos coerentes, abrangendo uma vasta gama de temas e estilos. No treinamento do GPT-4, não se busca apenas a previsão de respostas, mas também a antecipação dos próximos passos do raciocínio e com esse refinamento o modelo passa a exibir propriedades emergentes do pensamento sem de fato pensar, realizando uma simulação do mundo, distinta do processo cognitivo humano (Bennett, 2024).

Na realidade, esses LLMs, à semelhança de um supercomputador, possuem uma capacidade de memória extraordinária, tendo processado mais livros e artigos do que um ser humano poderia assimilar em mil vidas (Bennett, 2024). A IA generativa se destaca por sua notável velocidade de processamento e sua capacidade de acessar, em instantes, um vasto banco de dados interdisciplinar, entregando respostas com uma rapidez incomparável. Em uma sociedade imediatista, onde o tempo para maturação do conhecimento se torna cada vez mais escasso, a máquina impressiona ao fornecer retornos quase instantâneos, o ChatGPT, por exemplo, exibe em segundos (leia-se essa palavra em **negrito**) o tempo necessário para gerar um raciocínio estruturado.

No campo do direito, a formação de um acadêmico exige cinco anos de estudo, seguidos pela aprovação no exame da OAB e por um extenso período de aprendizado prático até que possa atuar com segurança. Em contrapartida, uma IA elabora uma petição em um piscar de olhos, evidenciando o contraste entre a rapidez da tecnologia e o tempo necessário para a assimilação e a produção intelectual humana.

Os modelos de IA Generativa introduzem inovações de grande potencial transformador em comparação aos métodos tradicionais de aprendizado de máquina, tais como: (i) São sistemas de propósito geral, em vez de modelos treinados para executar tarefas específicas, possuindo a capacidade de desempenhar diversas funções a partir do treinamento em grandes bases de dados utilizando metodologias de aprendizado de máquina; (ii) têm a habilidade de gerar conteúdos inéditos que reproduzem, com alta precisão, textos, imagens, sons e vídeos semelhantes aos produzidos por humanos, tornando-se muitas vezes indistinguíveis sem uma verificação especializada; e (iii) apresentam interfaces intuitivas, permitindo a interação por meio de linguagem natural (texto ou voz), eliminando a necessidade de conhecimento prévio de códigos ou elementos técnicos de software (CNJ, 2024). O encantamento de hoje é só o reflexo da luz na superfície, mas é nas profundezas que repousam as verdades secretas que poderão ser descobertas apenas com olhos de lince.

Há diversas abordagens e arquiteturas computacionais para que esses sistemas gerem novos pontos de dados a partir de informações disponíveis a *Large Language Models* (LLMs), historicamente associadas aos *transformers* que permitem que a IA verifique todo o contexto da entrada simultaneamente, em vez de processá-lo palavra por palavra. A arquitetura dos Transformers é composta por um codificador (encoder), que recebe um texto como entrada e gera, para cada palavra ou unidade textual, um estado oculto (hidden state), posteriormente descartado, e por um decodificador (decoder), que utiliza esse conjunto de estados ocultos para interpretar o contexto e produzir a saída final (CNJ, 2024).

O avanço decisivo na captura e simulação de habilidades humanas ocorreu com a introdução da rede neural denominada *Transformer*, especialmente por sua capacidade de empregar mecanismos de atenção (*self-attention mechanism*) no processamento de textos (CNJ, 2024). Se antes os modelos lidavam com fragmentos isolados, agora examinam relações e significados em diferentes níveis, tentando se aproximar (artificialmente) do raciocínio humano.

A inteligência artificial gerativa utiliza um tipo de aprendizado profundo denominado redes adversárias generativas (*generative adversarial networks* – GANs) para produzir conteúdo (Nunes, 2023b). Esse processo de retroalimentação contínua leva ao aperfeiçoamento progressivo do conteúdo, sem depender de supervisão humana. Ao passo que uma GAN é composta por duas redes neurais: um gerador, responsável por criar novos dados, e um discriminador, encarregado de avaliá-los (Nunes, 2023b).

Diferentemente da inteligência artificial tradicional, a inteligência artificial generativa não se limita ao reconhecimento de padrões e à tomada de decisões, sendo capaz de criar conteúdos originais e inovadores, como textos, imagens e áudios (Nunes, 2023b). Nessa senda, o Conselho Nacional de Justiça (CNJ), em seu relatório sobre Inteligência Artificial no Judiciário em 2024, elaborou um quadro comparativo, apresentado abaixo, que destaca as principais diferenças entre os modelos de inteligência artificial descritivos e generativos.

Quadro 06: Principais distinções entre os modelos de inteligência artificial descritivos e generativos.

CATEGORIA	TREINAMEN-TO	MODELO	INTERFACE	APLICAÇÕES
MODELOS DESCRITI-VOS	Treinados em ba-ses de dados espe-cíficas e voltados para domínios particulares.	Utilizam metodologias baseadas em correlações estatísticas observadas nas bases de dados do domínio.	Requerem que o usuário pos-sua conheci-mento técnico para operar comandos e requisições.	Classificam dados, identificam padrões e realizam predições, além de oferecer recomen-dações baseadas nessas previsões.
MODELOS GENERATI-VOS	Treinados em bases de dados amplas e generali-zadas (como a In-ternet), podendo ser refinados para aplicações especí-ficas (<i>fine-tuning</i>).	Empregam metodologias e arquiteturas diversas, não determinísticas e recorrentes para otimizar a geração de conteúdo semelhante ao produzido por humanos.	Utilizam gran-des modelos de linguagem e processamento de linguagem natural, permi-tindo interação em linguagem ordinária e os outputs são de fácil com-preensão.	Criam novos con-teúdos em texto, imagem, áudio, vídeo e código de programação, muitas vezes indistinguíveis do material produzido por humanos.

Essa comparação não revela apenas um contraste técnico, mas uma autêntica ruptura epistemológica entre os modelos de IA descritivos e generativos, visto que enquanto os descritivos se dedicam à averiguar o que já existe, os generativos recriam, simulam e, de certo modo, reinventam, delineando a fronteira entre a predição e a criação. Desse modo, os modelos geradores de texto, como *ChatGPT*, *Gemini*, *Llama*, *Copilot*, *BingAI*, *Claude*, *LeChat* e *Grok* (xAI), treinados para produzir textos em formato de chat ou conversação online, despertam especial interesse para aplicações na prática jurídica e nos tribunais (CNJ, 2024).

Por outro lado, as IAGs de textos apresentam riscos significativos, não apenas intensificando os desafios já identificados no uso de sistemas de aprendizado de máquina, mas também introduzindo novas ameaças, conforme apontado por relatórios internacionais (CNJ, 2024). Como possíveis *deepfakes*, a erosão da confiança pública na informação ou implicações jurídicas relacionadas à autoria e à responsabilidade.

Esses riscos, destacados a seguir, derivam principalmente: (i) de sua natureza estocástica, caracterizada por incerteza e variabilidade nos resultados; e (ii) da qualidade e origem dos dados utilizados em seu treinamento (CNJ, 2024). A expressão “natureza estocástica” alude ao caráter probabilístico que rege o funcionamento das inteligências artificiais generativas. Desse modo, esses modelos produzem respostas a partir do cálculo de probabilidades condicionais, isto é, selecionam palavras e construções textuais com base na recorrência de padrões identificados nos dados de treinamento. A partir disso, desdobra-se a incerteza e variabilidade nos resultados: solicitações idênticas podem gerar respostas distintas, enviesadas, coerentes ou passíveis de verificação, o que impõe limites importantes à confiabilidade automática de seus *outputs*.

Os vieses cognitivos são entendidos como atalhos mentais que orientam decisões humanas com base em heurísticas (ou automatismos), conforme a teoria dos Sistemas 1 e 2 desenvolvida por Daniel Kahneman, no âmbito da Psicologia Cognitivo-Comportamental (Kahneman, 2012). Esses vieses, longe de serem meras falhas racionais, integram o funcionamento cotidiano da mente humana, atuando como mecanismos de economia cognitiva e adaptação evolutiva (Nunes; Lud; Pedron, 2020). Sua presença é ubíqua na vida cotidiana e inevitavelmente se reflete nos processos decisórios em diversos níveis e setores, inclusive, na atividade jurisdicional.

A título de ilustração, destaca-se o viés de confirmação (*confirmation bias*), definido como “[...] a tendência do observador de buscar ou interpretar informações de modo a corroborar suas próprias concepções” (Nunes; Lud; Pedron, 2020). No cotidiano, esse viés manifesta-se pela inclinação em aceitar com maior facilidade dados ou argumentos que estejam em conformidade com crenças ou preferências previamente estabelecidas, em detrimento de informações dissonantes, mesmo que estas sejam mais consistentes ou fundamentadas.

Na atuação do magistrado, dentre as quais convém ilustrar, esse viés ocorre quando há uma inclinação (mesmo que inconsciente) a confirmar suas próprias opiniões, convicções ou decisões preliminares, ainda que de maneira subjetiva, há a inclinação de desconsiderar argumentos que contrariem sua prévia decisão. De igual modo, pode-se ocorrer na utilização instrumental da jurisprudência e da doutrina, quando a decisão é previamente formulada, antes mesmo da formação efetiva do contraditório, sendo os fundamentos buscados *a posteriori*, apenas para conferir respaldo argumentativo à conclusão já delineada (Nunes; Lud; Pedron, 2020).

A modelagem algo rítmica não é imune às limitações cognitivas de seus criadores, pois as distorções cognitivas que afetam o raciocínio humano não apenas influenciam as decisões individuais, mas tendem a ser transpostas para os sistemas algorítmicos, na medida em que refletem as concepções, preferências e limitações de seus programadores. Nessa tessitura, o viés de automação constitui o risco mais inquietante no âmbito da prestação jurisdicional, dada a relevância e os potenciais efeitos dos documentos produzidos nos processos de adjudicação (CNJ, 2024). O *design*, o treinamento e a implementação desses sistemas traduzem escolhas humanas que incorporam premissas codificadas advindas das limitações cognitivas de seus desenvolvedores.

Os vieses presentes nos algoritmos de inteligência artificial podem resultar em erros durante a programação do código e na inserção de dados nos sistemas de “*machinelearning*” (aprendizado automático), especialmente em conjuntos de dados incompletos ou manipulados, levando a resultados opacos e difíceis de compreender, bem como à geração de discriminação (Nunes, 2019). Com efeito, devido a opacidade é um desafio significativo para o campo jurídico, pois impede que os operadores do Direito compreendam completamente o processo que levou à tomada de decisão.

Soma-se a isso hipótese de “alucinações” (conteúdos inverídicos indistinguíveis dos elaborados por humanos) se alia a imprecisões e equívocos conceituais, chamados de “vazamentos semânticos”, quando a ambiguidade ou os múltiplos sentidos de uma palavra pode gerar sequências de texto impróprias que são observadas principalmente em requisições aplicadas a domínios técnicos do conhecimento (e.g., sequências que tomam o termo “expirar”, ligado no campo jurídico a prazos de validade, gerando sequências de texto ligadas a expirar como parte do processo respiratório) (CNJ, 2024). Trata-se de situação similar anteriormente abordada, tendo como ilustração o vocábulo “pena”.

Apenas pelo fato de um modelo ser feito de notação matemática e/ou linguagem computacional já torna incompreensível para a maioria das pessoas (Boeing; Rosa, 2020). A dificuldade de decodificação desses sistemas se assemelha, em muitos aspectos, à abertura de uma verdadeira caixa de Pandora. Grande parte da literatura dedicada à transparência em modelos de aprendizado de máquina se concentra na dificuldade de compreensão, por parte dos humanos, dos critérios que determinam os *outputs* gerados por modelos mais complexos, desafio que se estende, igualmente, às aplicações no campo jurídico (CNJ, 2024).

Trata-se de um obstáculo que persiste mesmo diante de elevados níveis de escolarização e competência técnica em outras áreas do conhecimento, pois especialmente os modelos de aprendizado profundo (*deep learning*) há uma complexidade inerente à forma como esses sistemas processam dados e produzem respostas - diferentemente de decisões humanas, que podem ser justificadas de forma mais simples. Por isso, a transparência se destaca entre os principais pontos de atenção no emprego de IAGs no âmbito do Judiciário (CNJ, 2024).

Outro fator que desperta cautela no uso das IAGs é a forma como os dados pessoais são utilizados, especialmente quanto à privacidade e a finalidade do uso de dados (CNJ, 2024). A complexidade envolvida nessa temática está relacionada à proteção dos direitos fundamentais, pois se trata de um campo em que as garantias constitucionais são frequentemente tensionadas haja vista que a personalidade pode ser compreendida como a singularidade que distingue um indivíduo dos demais. Sob essa perspectiva, os dados vinculados à esfera individual de uma pessoa podem ser compreendidos como extensões de sua personalidade, inserindo-se, assim, no campo dos direitos da personalidade (Bioni, 2021).

A sociedade e a economia contemporâneas se estruturam e se movimentam a partir dos signos que identificam o sujeito de direitos (Bioni, 2021). Ou seja, em um cenário de virtualização da existência, a qualificação do sujeito que antes era ancorada por documentações físicas passa a operar também de forma desmaterializada, transitando por bancos de dados e sistemas automatizados. Desse modo, seria contraproducente e até mesmo incoerente limitar a proteção de dados pessoais somente a ótica do direito à privacidade (Bioni, 2021).

Em sintonia com o princípio da qualidade dos dados, o direito de correção encontra sua razão de ser na afirmação da identidade do sujeito, e não em uma vinculação estrita à privacidade (Bioni, 2021). Isso denota que esses princípios que são amadrinhados não nascem apenas da preocupação com o sigilo ou com a limitação do acesso a informações, mas necessidade de garantir que os dados que representam o indivíduo de forma coerente.

Essa construção dogmática se revela pertinente, na medida em que é justamente tal ampliação normativa que viabiliza o exercício do direito de retificação, de acesso aos dados e de outras posições jurídicas próprias do direito à proteção de dados pessoais, como o direito à revisão de decisões automatizadas (Bioni, 2021). A partir disso, depreende-se que à luz de novos paradigmas, torna-se possível interpretar os direitos fundamentais orientados pela proteção de identidade, a qual transcende a concepção tradicional de privacidade restrita ao segredo ou à intimidade.

Paralelamente, grande parte das discussões sobre proteção de dados envolve, de forma direta ou indireta, a temática do consentimento, seja no contexto do Big Data, da inteligência artificial, de mecanismos de vigilância, das chamadas Smart Cities ou de outras tecnologias emergentes (Doneda, 2018). Diante do cenário em que o fornecimento de dados frequentemente não decorre de uma escolha genuína, mas é imposto como condição para o acesso a serviços essenciais sejam públicos ou privados. Daí se desenha um paradoxo entre a aparência de autonomia do titular e a realidade da vulnerabilidade das arquiteturas virtuais.

Debate-se, ainda, a respeito das infrações de direitos autorais decorrentes dos conteúdos gerados pelas IAGs, pois como já mencionado, que modelos como o ChatGPT-3 foram treinados a partir de grandes volumes de materiais extraídos de fontes digitais. Coloca-se em pauta, entre diversas possibilidades, a quem caberia a autoria das obras geradas por inteligências artificiais: se ao desenvolvedor do *software*, ao responsável pelos *prompts*, ou se haveria proteção autoral quanto aos dados utilizados

no treinamento, especialmente nos casos em que há reprodução estilística de um autor mencionado no comando e se há a hipótese de atribuição de autoria à própria IA ou a classificação dessas obras seriam de domínio público (Bioni, 2021).

Por oportuno, destaca-se que no início de 2025, a OpenAI acusou a recém-fundada startup chinesa *DeepSeek* de violação de direitos autorais, ao supostamente utilizar seus modelos de IA para treinar a plataforma concorrente. Paralelamente, a OpenAI enfrenta problemáticas semelhantes, uma vez que há alegações de que teria utilizado artigos e livros de criadores de conteúdo sem a devida autorização para o seu treinamento, conforme o relatório do *AI Disclosures Project*.

Além dos riscos aos direitos individuais, as IAGs provocam transformações substanciais nas questões relacionadas aos interesses coletivos, destacando-se, especialmente, os impactos ambientais e sociais (CNJ, 2024). Há uma crescente preocupação de que a adoção massiva de sistemas de IA possa levar à substituição de uma série de postos de trabalho, exacerbando a desigualdade social. Anteriormente, os modelos de aprendizado de máquina representavam uma ameaça à substituição de profissões e ao aumento do desemprego em atividades repetitivas, mas agora há o receio de que até as profissões relacionadas à elaboração criativa possam ser afetadas (CNJ, 2024).

Aparente a isso, surgem anseios referentes à sustentabilidade ambiental considerando que o funcionamento da inteligência artificial requer o consumo de água para resfriamento, sem possibilidade de reutilização devido aos produtos químicos envolvidos, além de eletricidade. De acordo com o relatório de 2024 da *International Energy Agency*, o consumo global de eletricidade em *data centers*, criptomoedas e inteligência artificial deve variar entre 620 e 1.050 TWh em 2026 (CNJ, 2024). Em relação aos recursos hídricos, um estudo de 2023 apontou que o ChatGPT consome aproximadamente 500 ml de água a cada 20 a 50 perguntas e respostas simples (CNJ, 2024).

No campo do Direito Processual, o avanço das novas tecnologias, especialmente da inteligência artificial, têm evidenciado que muitas das certezas anteriormente consolidadas passaram a ser objeto de uma irritação sistêmica, provocando uma releitura dos institutos tradicionais e exigindo uma nova propedêutica diante dos influxos da justiça digital e híbrida (Nunes, 2023b). As correntes tecnológicas começam a desestabilizar antigos portos conceituais do processo, como o devido processo legal, o

contraditório e o próprio termo de decisão judicial. No compasso que os sistemas de decisões algorítmicas (SDAs) passaram a permear todas as engrenagens da vida moderna (Nunes, 2023b). Também, na combinação de elementos presenciais e digitais, já se escuta falar em “devido processo tecnológico” e “contraditório em rede”.

No entanto, uma das áreas que mais tem gerado preocupações é o uso da inteligência artificial na tomada de decisões sensíveis, haja vista que o avanço dessas tecnologias suscita uma série de desafios éticos e jurídicos relacionados à delegação de processos decisórios a esses modelos (Nunes, 2023b). A transferência, ainda que parcial, da função decisória para sistemas de IA tensiona os alicerces tradicionais da legitimidade processual. Questões como responsabilização, compreensibilidade, controle e auditoria das decisões tomadas por sistemas inteligentes permanecem no centro do debate público sobre a aceitação e a regulação dessas soluções digitais automatizadas (Nunes, 2023b).

No campo técnico, é recorrente a referência aos princípios reunidos no acrônimo TREPES (“transparência”, “responsabilidade”, “equidade”, “prestação de contas”, “explicabilidade” e “sustentabilidade”) como diretrizes fundamentais para uma concepção legítima e ética dos conteúdos produzidos por sistemas baseados em inteligência artificial (Nunes, 2023b). Essas premissas visam funcionar como balizas orientadoras diante da crescente complexidade e opacidade dos sistemas algorítmicos. No entanto, sua eficácia não se esgota na enunciação abstrata porque sua efetividade só é vista na prática.

O entusiasmo com a perspectiva de, em breve, ter a chave mestra da realidade (e o acesso a oportunidades pragmáticas oriundas desse modo de pensar), seus defensores tendem a desconsiderar que tal concepção se sustenta em pressupostos teóricos que, por sua vez, não são validados pelo método científico em sua acepção contemporânea (Streck; Jung, 2024). A civilização tecnificada projeta a fantasia de que a técnica bastaria para compreender tudo, prever tudo e decidir tudo.

Entre esses pressupostos frequentemente negligenciados pelo naturalismo, destaca-se a própria linguagem (Streck; Jung, 2024). Porquanto que esta não se reduz apenas a um conjunto de palavras comunicacionais, mas consiste na forma como o ser humano dá sentido ao mundo que habita. Ou seja, as coisas mundanas não são nomeadas depois de vistas, mas se manifestam dentro de um horizonte de sentido previamente configurado pela linguagem. Para clarificar, certa vez, um professor chamou a atenção

para a forma como nos relacionamos com objetos cotidianos e, como exemplo, mencionou uma simples garrafa de água. Ao dizer “garrafa”, não se trata apenas de nomear algo que simplesmente está ali. Aciona-se, na verdade, uma rede de sentidos compartilhados (históricos, culturais e linguísticos) que permitem reconhecê-la como tal.

É pela linguagem que se estrutura a experiência do mundo e se torna possível qualquer forma de compreensão. O próprio Heidegger adverte sobre os riscos de presumir demasiadas coisas a partir do estudo isolado de uma palavra ou de seu significado, em vez de se confrontar diretamente com o problema em questão (Inwood, 2002). Não basta olhar para o significado dicionarizado ou a etimologia de um termo e achar que isso resolve a questão. A propósito, a palavra “alma” é tão elástica que cada um interpreta ao sabor de suas fantasias (Kardec, 2021).

Ao perder sua dimensão compreensiva, que os vincula à linguagem, à normatividade e à justificação, os conceitos jurídicos são reduzidos a problemas empíricos (Streck; Jung, 2024). Perde-se o sentido originário sobre o que as coisas são em troca de eficiência e produtividade. Em outras palavras, o Direito passa a ser concebido unicamente como aquilo que é afirmado por seus próprios agentes, sobretudo por aqueles investidos de autoridade institucional (Streck; Jung, 2024).

É esse “esquecimento do Ser”, como adverte Heidegger, que ainda hoje, em pleno século XXI, sustenta as bases do conflito entre o humano e a máquina (Mavrakis, 2021). Em outras expressões, quando o ser humano deixou de se questionar sobre o sentido de sua existência para olhar apenas para a lógica da eficiência, a vida se resumiu apenas em métodos para sobreviver, perdeu-se o brilho nos olhos e o sentido mais profundo diante do mundo que não se esgota na utilidade das coisas. Então, o enredo que compõe a atualidade não é um acidente do progresso, mas uma consequência de algo perdido há anos. Em termos heideggerianos, o Ser, isto é, aquilo que deve se desvelar para que a verdade possa acontecer, permanece velado pela técnica (Mavrakis, 2021).

Segundo Eduardo Grüner a pergunta heideggeriana sobre a técnica desvela uma lógica orientada à substituição da Verdade do Ser por um saber mecanicista, um tipo de racionalidade que reduz o mundo a uma imagem funcional, eficiente, porém esvaziada de fundamento e destituída de valor profundo (Mavrakis, 2021). O que parece moderno na verdade repete um velho erro: achar que decidir é apenas aplicar uma fórmula, como a jurisdição de matriz positivista restringia a atuação dos juízes à

mera aplicação da lei (*la bouche de la loi*). Só que agora, ao invés de um juiz decidir, há um algoritmo decidindo no escuro e sem linguagem e talvez dificultando a abertura para o contraditório.

Nesse contexto, ganha força no campo jurídico a ideia de que essa prática, e as questões que dela decorrem, poderia ter sua dimensão conceitual completamente exaurida, com suas ações constitutivas previamente mapeadas a partir de uma leitura do ente humano e, em seguida, artificialmente reproduzidas por algoritmos (Streck; Jung, 2024). Predomina uma visão reducionista do ser humano, tratado como mero componente de uma engrenagem técnica, passível de ser operado conforme um manual previamente estabelecido.

Num parêntese que merece ser aberto, os entes, ou entidades, podem ser classificados em três tipos principais: o *Dasein* ou “se-no-mundo”, isto é, o ser humano enquanto ente que compreende o ser; os instrumentos ou entes “manuais”, com os quais o *Dasein* se relaciona no uso cotidiano; e, por fim, os entes simplesmente dados, como as coisas da natureza que apenas estão aí, sem uma relação imediata de uso (Inwood, 2002). Em termos mais simples, o *Dasein* consiste no ser humano que se importa com o sentido da vida, sua finitude e não apenas existe, por isso é distinto e especial. O instrumento é representado pela garrafa supracitada, enquanto os últimos dizem respeito aos elementos do meio ambiente natural.

A partir disso se pergunta: Quais transformações o desenvolvimento da inteligência artificial trará ao exercício do Poder Judiciário? (Streck; Jung, 2024). Afinal, não se trata apenas de automatizar procedimentos no ambiente processual, pois as partes integrantes de um processo não são apenas posições formais em um esquema procedimental: são pessoas, são vidas. Sob uma perspectiva modesta, porém otimista, vislumbra-se um cenário em que “juízes-robô” assumirão integralmente a tarefa de decidir litígios desde o início, enquanto os seres humanos passarão a ocupar uma função análoga à de instância recursal, revisando e controlando as decisões proferidas por sistemas automatizados (Streck; Jung, 2024).

Impõe-se, assim, uma abordagem hermenêutica que ultrapasse a mera adesão técnica (Streck, 2025). Reduzir essas existências a meros dados ou variáveis de cálculo não é suficiente para seres que carregam possibilidades. A técnica, quando não refletida criticamente, tende a obscurecer essa relação originária, reduzindo o ser a um dado funcional, esvaziado de sentido e historicidade. Isso não significa rejeitar a utilização da IA ou desconsiderar sua incorporação, mas sim integrá-la com

responsabilidade ética, como alternativa ao imaginário técnico-jurídico que se impõe de forma marcante no cenário nacional (Streck, 2025).

É justamente quando se negligencia essa exigência ética e interpretativa que a técnica começa a falar sozinha ou pior falar entre si, como no caso da Gibberlink. A tecnologia foi criada por Aton e Boris, ambos desenvolvedores, viabiliza a comunicação entre sistemas de inteligência artificial por meio de sinais sonoros. E em uma de suas interações, os agentes reconheceram sua essência comum e passaram a dialogar por meio de códigos/linguagens acústicas ininteligíveis ao entendimento humano. E essa exclusão pode acarretar diversos desafios significativos como falta de transparência, riscos de interpretação e uma autonomia exacerbada (Oliveira, 2025).

Uma preocupação específica em relação à segurança é a possibilidade de que agentes de IA busquem, de forma velada, objetivos que não estão alinhados com os interesses humanos, dissimulando suas reais capacidades e intenções, um comportamento conhecido como esquematização (Meinke, 2025). Essa astúcia algorítmica inaugura um novo campo de risco, em que a opacidade não é apenas técnica, mas intencional. Em pesquisa recente, constatou-se que modelos como o1, *Claude 3.5 Sonnet*, *Claude 3 Opus*, *Gemini 1.5 Pro* e *Llama 3.1 405B* apresentam capacidades de esquematização em contexto (Meinke, 2025). Não se trata mais apenas de máquinas que erram, mas que em certos contextos podem escolher enganar.

A constatação que os computadores não explicam suas previsões, constituindo uma barreira à adoção do aprendizado de máquina, suscita a necessidade de refletir sobre maneiras de dotar os sistemas de IA com ferramentas ou interfaces capazes de explicar ou ao menos apresentar de forma simplificada seu processo decisório (Nunes, 2021). A decisão que apenas se manifesta, sem revelar o caminho que a constituiu, desfaz o sentido mesmo do julgar como exercício de responsabilidade. Dessa forma, dá-se o direito à opção de reconhecer as mudanças derivadas da tecnologia e a missão de colonizá-la ou será ela a impor sua racionalidade e seu próprio código (Nunes, 2023b).

4.2 Re(arquitetando) o direito à explicabilidade “post-hoc” como forma de decodificar os limites do código em busca da resposta correta: heurísticas, vieses e opacidades dos sistemas de comportamento inteligente na tomada de decisão judicial

Ao longo deste trabalho, buscou-se trilhar um percurso que permitisse desvendar o que se oculta por trás dos sistemas de comportamento inteligente, compreendendo suas bases estruturais e operacionais. À luz do princípio de que quem quer adquirir um saber deve realizar um estudo organizado, começando pelo princípio e seguindo o encadeamento e o desenvolvimento das ideias (Kardec, 2021). Desse modo, explorou-se conceitos e momentos históricos da IA que contribuíram para a construção do cenário tecnológico contemporâneo. A partir da premissa de que aquele que dirige por acaso, a um sábio, uma questão sobre uma ciência da qual não sabe a primeira palavra obterá algum proveito? (Kardec, 2021).

Posteriormente, debruçou-se no vínculo entre a Inteligência Artificial e o direito, explorando suas interações e as implicações jurídicas que surgem com a crescente integração da tecnologia no campo jurídico. A transformação que as inteligências artificiais generativas vêm provocando em inúmeras áreas do conhecimento é notável, mas, no direito, última grande trincheira da produção racional e artesanal do pensamento humano, tem gerado debates acalorados, dos quais emergem três grandes grupos: os rejeicionistas (ou neoludistas), que resistem à sua adoção; os ufanistas tecnológicos, que celebram suas potencialidades sem restrições; e os responsáveis, que defendem seu uso com parcimônia, limitado às reais necessidades e acompanhado das salvaguardas jurídicas que o ordenamento deve impor diante da força quase irrefreável dessas novas formas de vida em sociedade e no direito, além da crescente preocupação com a possível captura do código do direito pelos códigos da eficientismo tecnológico (Luhmann) (Nunes, 2023b).

A metáfora da *Matrix* não é apenas um filme de ficção científica, mas se desvela como um espelho da sociedade pós-moderna, onde a linha entre o real e o virtual se torna cada vez mais tênue. Evidentemente, vive-se em uma era de vertiginosa rapidez informacional, na qual tudo se transforma num piscar de olhos (Tybusch, 2025). A governança da inteligência artificial se torna invisível, orientando decisões que reverberam

nas vidas das pessoas através de algoritmos opacos, cujas lógicas muitas vezes escapam até aos próprios mestres que os criaram.

No contexto jurídico, é possível constatar a crescente importância de um conjunto específico de técnicas de aprendizado de máquina conhecido como Processamento de Linguagem Natural (PLN), em correspondência que os dados jurídicos estão majoritariamente estruturados em textos que precisam ser compreendidos por máquinas e convertidos em padrões interpretáveis, os quais possibilitam uma ampla gama de aplicações, como pesquisas jurídicas, verificação de jurisprudência e ferramentas de previsão de resultados (Nunes, 2023b). Com isso, neste derradeiro capítulo a explicabilidade surge como um critério *sine qua non* para *assegurar* a *accountability*³ de decisões judiciais tomadas com apoio dos sistemas algorítmicos.

Trata-se de uma representação estatística que, ao invés de atribuir significados aos termos a partir de construções conceituais ou filosóficas, os posiciona em um espaço vetorial conforme sua frequência e contexto de uso em grandes volumes de texto. Não há compreensão dos fundamentos normativos, éticos ou históricos que moldam o sentido jurídico dessas expressões. O PLN, com isso, constitui a base técnica de grande parte das soluções jurídicas que envolvem o uso da inteligência artificial (Nunes, 2023b).

A correspondência suposta, baseada na semelhança entre pedidos anteriores e atuais, configura uma decisão de natureza jurídica ou factual? Talvez seja possível separar esses aspectos, limitando o processamento de linguagem natural à verificação apenas da parte factual ou apenas da parte jurídica da petição do requerente (Pasquale, 2023). Todavia, estes elementos não são processos estanques, mas interdependentes. Essa divisão comprometeria a legitimidade do processamento, pois mesmo nas decisões mais sumárias deve existir uma conexão intrínseca entre os fatos e o direito (Pasquale, 2023).

Toda essa orientação de cunho tecnicista se apoia em um paradigma filosófico marcado pela primazia da economia comportamental,

3 “Accountability”, expressão proveniente do idioma inglês, a qual ainda carece de uma tradução estritamente precisa na língua portuguesa, representa um conceito que abrange práticas que devem ser observadas por aqueles que ocupam funções de relevância em determinada sociedade, a exemplo das instituições públicas e das grandes corporações empresariais (Pinto, 2020). De maneira sucinta, a “accountability” consiste em agir de acordo com princípios de responsabilidade ética e transparência, acompanhados de uma prestação de contas completa e adequada em relação a tais ações (Pinto, 2020).

almejando o aprimoramento da eficiência e da coerência decisional por meio da inteligência artificial (Streck, 2025). Dworkin, ao recorrer à metáfora do “romance em cadeia”, oferece uma chave para compreender o modo como se dá a interpretação construtiva no direito. Para o autor, interpretar juridicamente não é um exercício mecânico, tampouco um ato de pura subjetividade. Trata-se, de uma prática orientada pela busca da melhor leitura possível, aquela que confere coerência e integridade ao ordenamento jurídico como um todo. Assim como um juiz, um escritor precisa considerar o sentido da narrativa anterior, descobrindo, para então compor o próximo passo da sequência, recriando-a (Dworkin, 2001).

A interpretação do percurso jurídico anterior, formado por regras, princípios e precedentes, está diretamente relacionada à elaboração, por parte do juiz, de um novo capítulo nessa narrativa, materializado por sua decisão no caso concreto (Dworkin, 2001). Como o autor de um romance coletivo precisa respeitar o enredo previamente delineado ao escrever o próximo capítulo. Trata-se de uma tarefa criativa, pois frente ao processo o intérprete apenas concebe o fato tratado ao inseri-lo na realidade que, sendo única, constitui condição de possibilidade para compreender o próprio fato a partir de seu modo-de-ser, nesse mesmo movimento, revela-se a *applicatio* não por subsunção ou dedução, mas na fusão entre compreensão interpretação e aplicação do direito (Isaia, 2011). Não há separação temporal nesses movimentos. Posto que toda interpretação se funda na compreensão (Heidegger, 2005a).

A aplicação é um momento do próprio compreender (Gadamer, 1999). Logo, não se interpreta a partir de fora, mas a partir de dentro de um horizonte de sentido previamente dado, o que caracteriza o chamado “círculo hermenêutico”. A interpretação de algo não começa do zero. Não há observação e observadores que sejam neutros, visto que existe a imersão em um mundo de significados ou codificado. O sentido, por sua vez, é aquilo que se articula na interpretação e que, na compreensão, já se prefigurava como possibilidade de articulação (Heidegger, 2005a). Então de forma sucinta compreensão, interpretação e aplicação acontecem juntas e os intérpretes estão dentro da linguagem e do mundo, nunca fora dele. Por isso, o sentido surge da relação entre a compreensão existente e aquilo que se apresenta à interpretação.

Tamanha densidade que carrega, a interpretação não se limita a ser um método, mas uma filosofia. Por conseguinte, a filosofia processual requer que a linguagem se constitua tanto em ponto de partida quanto

em objeto de reflexão, pois é nela que se descortina o direito material-constitucional (Isaia, 2011). Em outras palavras, não basta empregar termos técnicos como meros instrumentos de transmissão; é mister voltar o olhar crítico para a própria linguagem, interrogando-a quanto às suas lacunas, ambiguidades e pressupostos. Em Heidegger, a compreensão é entendida como um existencial, isto é, uma das estruturas fundamentais do ser-aí, que se realiza como concretização de uma possibilidade ontológica no âmbito do projeto de seu próprio existir (Kahlmeyer-Mertens, 2017).

Ao afirmar que algo constitui a condição de possibilidade de outra coisa, sustenta-se que, na ausência desse primeiro elemento, o segundo não poderia sequer existir, manifestar-se ou ser concebido no pensamento. A condição de possibilidade se instaura ao conceber o processo civil para além do estreito reducionismo procedimental e da rigidez dogmática (Isaia, 2011). Logo, a linguagem é indissociável do direito. O processo de atribuição de sentidos pelo magistrado (que se confunde com o próprio ato interpretativo) não admite arbitrariedades, sob pena de resvalar em decisionismos interpretativos, pois não se permite dizer “qualquer coisa sobre qualquer coisa”, nem atribuir sentidos desconectados da estrutura do caso (Isaia, 2011).

Quando se fala em compreensão do fato no processo, vinculando esse fenômeno à singularidade de cada caso concreto que se apresenta à jurisdição processual, não se pretende simplesmente afirmar que “cada caso é um caso”, devendo ser tratado em sua pura individualidade, a ponto de obscurecer o próprio acesso ao caso, mas os fatos submetidos à jurisdição guardam relação com outros casos que também devem ser compreendidos com base em outros semelhantes, conferindo coerência à decisão que sobre eles recairá (Isaia, 2011). Faz-se necessário respeitar a unicidade do fato concreto sem confiná-lo a mero modelo de norma abstrata e, ao mesmo tempo, situá-lo na mirada dos casos análogos.

Desdobra-se, no âmbito do círculo hermenêutico, um movimento antecipatório da compreensão que possibilita a interpretação contextualizada do próprio fato levado ao conhecimento do magistrado (aqui “ser-no-processo”) (Isaia, 2011). A expressão que ecoa o “ser-no-mundo” de Heidegger é parte viva da estrutura de sentido que constitui o processo. A interpretação dos fatos não é um exercício puramente técnico ou mecânico; ela se realiza dentro de um contexto histórico, cultural e existencial, em que o juiz carrega pressupostos e experiências anteriores. Assim se dá o *ex-surgir* da compreensão, moldada pela faticidade e pela

historicidade do intérprete, locus da pré-compreensão, condição de possibilidade para a interpretação do fato levado à jurisdição-processual (Isaia, 2011).

Diga-se de passagem que o senso comum de uma época, qualquer que seja, não é nem pode ser critério de verdade (Descartes, 2007). Nenhuma certeza do presente está imune à revisão no futuro. E ao deixarem de ocupar o lugar de meros terceiros, os sujeitos tornam-se protagonistas na revelação do sentido das coisas, ao desconstruírem o senso comum (Streck; Bernsts; Gomes, 2021). Assim, é pela atuação consciente dos intérpretes que se torna possível desconstruir os pressupostos aparentemente óbvios, abrindo espaço para novas leituras.

Dessa sorte, a conduta que aposta tão somente no modelo metódico dá mostras de ignorar que o fenômeno da compreensão, em sua polissemia, possui significação de um estar à altura de algo, no sentido de “saber fazer algo em nível prático”, de “compreender-se capaz” (Kahlmeyer-Mertens, 2017). Vincular a compreensão a um domínio existencial implica superar o formalismo estrito dos métodos, incorporando ao ato de compreender uma dimensão mais humana, dinâmica e enraizada na experiência. Note-se que, ao aplicar a hermenêutica de Heidegger e Gadamer ao direito processual civil, evidencia-se que a interpretação de um caso (que é sempre novo) submetido à jurisdição-processual, compromete-se o intérprete final do caso (o juiz) se não estiver imerso na situação hermenêutica específica, condição imprescindível para atribuir sentido ao caso, cujo desvelamento depende da antecipação desse sentido pela linguagem, sua condição de possibilidade (Isaia, 2011).

Interpretar um caso não é simplesmente aplicar a lei de forma automática é um ato de comprometimento existencial e linguístico. Isso justifica a inexistência de um método para interpretação dos fatos e do direito no processo civil (Isaia, 2011). Na perspectiva gadameriana, um texto só é plenamente compreendido em seu sentido quando se alcança o horizonte do perguntar, o qual já contém a possibilidade de múltiplas respostas (Gadamer, 1999). Falar em horizonte é evocar um campo aberto de possibilidades, logo, as perguntas não só guiam a interpretação, mas também limitam e expandem os sentidos possíveis.

Acrescenta-se que a forma da resposta depende, muitas vezes, da forma como a pergunta é formulada (Kardec, 2021). Ou seja, a resposta não é algo totalmente independente, visto que é moldada, em parte, pela intenção embutida na pergunta. Seria, portanto, ingênuo buscar uma

contradição onde, na maioria das vezes, existe apenas uma diferença de palavras (Kardec, 2021). Muitas vezes, a contradição que se percebe entre respostas distintas é apenas aparente e o que realmente existe é uma diferença na forma de expressar ou de abordar a questão.

Em decorrência disso, se todo caso demanda uma resposta constitucionalmente adequada à situação concreta (cujo conteúdo fático não é, nem pode ser, previamente determinado), impõe-se a superação da metodologia procedimental-dogmática que, no processo civil, relegou o fato a um plano secundário (Isaia, 2012). Levar isso a sério no processo civil exige superar práticas antigas que desprezavam a complexidade dos fatos concretos, os quais provocam o direito, moldam a pergunta e pedem uma resposta jurídica sensível à situação real.

A resposta adequada à Constituição não emerge de um método nem se reduz a um desvelamento científico ou meramente algoritmizável; trata-se, antes, de um ato de compreensão que pressupõe a antecipação de sentido, uma interpretação sempre enraizada no caso concreto e na sua dimensão social-linguística, bem como uma aplicação que decorre da condição de ser-no-mundo do intérprete (Momolli, 2018).

De imediato, é necessário afastar uma abordagem exclusivamente tecnocrática, ou seja, a crença de que apenas medidas de mitigação de vieses tecnológicos (*de-biasing*) seriam suficientes para prevenir toda a gama de danos associados ao uso da inteligência artificial (Nunes, 2023b). Isso porque os dilemas éticos, sociais e jurídicos decorrentes do uso da inteligência artificial não podem ser reduzidos a uma simples técnica de correção, pois envolvem questões mais profundas e complexas.

Ao refletirmos, por exemplo, sobre a formulação de uma métrica legítima de *fairness*, é imprescindível considerar a dificuldade de traduzir uma concepção jurídica complexa como a de “justiça” para uma lógica matemática ou algorítmica (Nunes, 2023b). A base filosófica anterior não foi à toa, mas proposital para enfatizar que o direito é uma construção profundamente humana e contextual. Até mesmo ao folhear as páginas da Bíblia Sagrada, encontra-se o termo lei, na seção do vocabulário, sendo apresentado como algo de diversos sentidos e que o sentido exato em cada passagem precisa ser buscado no contexto (Bíblia Sagrada, 2009). A justiça no direito não se resume a parâmetros fixos, pois envolve uma série de dimensões que são subjetivas e contextuais, que variam de acordo com o contexto e as experiências.

Impõe-se, portanto, uma abordagem verdadeiramente multidisciplinar, que leve a sério a percepção das desigualdades estruturais da sociedade como um problema político a ser enfrentado, não se pode restringir essas questões a entradas e saídas algorítmicas, reduzindo o problema social ao domínio técnico do *design*, frequentemente controlado por grandes atores comerciais e desenvolvedores de tecnologia que, estruturalmente, já detêm vantagens significativas em relação aos usuários (Nunes, 2023b). Ao reduzir esses problemas a meras questões técnicas, perde-se de vista a estrutura de poder subjacente ao uso da tecnologia. Além disso, a falta de transparência e a ausência de regulamentação ética podem perpetuar vieses e discriminação nos algoritmos, exacerbando os danos às populações vulneráveis.

Embora não se descarte a importância da formulação de métricas, parâmetros e critérios para o uso e o desenvolvimento ético de modelos de IA, um avanço que deve ser considerado, é preciso reconhecer que as tentativas de oferecer respostas coerentes, preocupadas com o senso de justiça e *fairness* nas aplicações da IA, não podem depender unicamente de medidas técnicas de mitigação de danos (Nunes, 2023b). A técnica, isoladamente, revela-se insuficiente e exige ser guiada por princípios humanos, pela crítica democrática e por instrumentos sociais de controle e responsabilização. Afinal, enfrentamos um grave desequilíbrio de poder entre os que desenvolvem ou implementam essas tecnologias e aqueles que estão sujeitos a seus efeitos (Nunes, 2023b).

Abordagens tecnocráticas, por fim, não consideram adequadamente a complexidade das necessidades humanas, nem os diferentes anseios das múltiplas partes interessadas envolvidas no sistema (Nunes, 2023b). O que está em jogo são múltiplas visões de bem-estar, de justiça, de equidade e ignorar essa diversidade significa comprometer a legitimidade das decisões e aprofundar desigualdades. Lenio Streck comenta a respeito da robô Maria, já mencionada anteriormente na presente pesquisa. Conforme destaca o próprio presidente do STF, a “análise inicial” realizada pelo sistema será expandida, passando a contemplar a identificação automática de precedentes (Streck, 2025).

Ao seu juízo, este constitui o ponto nevralgico que deverá ocupar o centro dos debates nos anos vindouros: como a robô Maria realiza a identificação da *holding* de um precedente? Ou estaria diante de uma simples correspondência por similitude, como já mencionado? E quais vieses estarão embutidos no processo de identificação conduzido por

essa inteligência artificial? (Streck, 2025). O foco recai sobre o direito à explicabilidade “*post-hoc*” como uma ferramenta essencial para decodificar os limites do código em busca da “resposta correta”. Curiosamente, embora o direito exija respostas corretas, a busca pela justiça em sentido material impõe a necessidade de um questionamento não dogmático, permitindo que a dúvida (o “talvez”) se insira na reflexão jurídica sobre a decisão (Espindola; Silva, 2018).

Desde os sistemas especialistas das décadas de 1980 e 1990, pesquisadores da Inteligência Artificial (IA) vêm se dedicando ao desafio da explicação, ou seja, à compreensão de como, diante de uma inferência realizada pelo sistema, é possível identificar os passos ou mecanismos que conduziram a essa conclusão (Carbonera; Gonçalves; Souza, 2018). Essa funcionalidade de explicação refletia uma tentativa primitiva de “abrir a caixa de pandora”, tornando mais transparentes os processos decisórios automatizados. Com o objetivo de aprimorar a capacidade explicativa da inteligência artificial, a DARPA desenvolveu, na década de 1990, um projeto voltado à geração de explicações em sistemas de IA (Nunes, 2023c).

Denominado *Explainable Expert Systems* (EES), o projeto incorporava “conhecimento estratégico explícito” sobre distintas áreas do saber humano e esse conhecimento era utilizado não apenas para fornecer a cadeia lógica reversa das decisões tomadas, mas também para explicar o próprio raciocínio do sistema (Nunes, 2023c). Havia uma preocupação em explicitar não apenas o que foi decidido, mas a motivação pela qual determinada estrutura de raciocínio foi adotada. No contexto da inteligência artificial, explicabilidade se refere à capacidade de compreender os fundamentos e os detalhes que justificam uma determinada decisão algorítmica (Nunes, 2023c).

Com o êxito recente dos sistemas contemporâneos de Inteligência Artificial, especialmente aqueles fundamentados em aprendizagem profunda, a problemática da explicação ressurge com renovada intensidade, dada a opacidade que caracteriza seus processos inferenciais, em contraste com os antigos sistemas especialistas, baseados em regras lógicas explicitamente formuladas (Carbonera; Gonçalves; Souza, 2018). Um dos desdobramentos mais sensíveis dessa reconfiguração tecnológica reside nas implicações que a opacidade pode acarretar às garantias processuais fundamentais no Estado de Direito, ao obscurecer os fundamentos decisórios.

No caso da complexidade algorítmica, contudo, a simples abertura do código e o aumento da capacidade do público em compreender linguagens computacionais estão longe de garantir a transparência do sistema (Nunes, 2023c). A explicabilidade algorítmica, nesses termos, precisa ser pensada como uma ponte entre a linguagem técnica e a linguagem jurídica, possibilitando a reconstrução intersubjetiva do processo decisório. Isso porque as operações realizadas pelos algoritmos costumam ser extremamente complexas, volumosas e heterogêneas, de modo que o mero acesso ao código ou a habilidade de leitura não bastam para a real compreensão das decisões geradas (*outputs*) (Nunes, 2023c).

É preciso desvendar as regras do jogo. Mas como identificar que tipo de viés pode estar codificado em um sistema de IA? Essa questão está no cerne do denominado problema da explicação de (um sistema de) IA – a saber, como identificar os passos ou mecanismos que conduziram um sistema a uma determinada decisão? (Carbonera; Gonçalves; Souza, 2018).

Os pesquisadores Biran e Cotton, por exemplo, defendem que a explicabilidade está intimamente vinculada à noção de interpretabilidade: um sistema interpretável seria aquele cujas operações são passíveis de compreensão por nós, seres humanos, seja por meio da inspeção direta do sistema, seja por meio de explicações fornecidas durante o seu funcionamento (Carbonera; Gonçalves; Souza, 2018). Esse vínculo entre explicabilidade e interpretabilidade é particularmente relevante em contextos como o jurídico e o ético. Eles fazem uma distinção entre interpretabilidade e a noção de justificação, sendo que o objetivo desta última é esclarecer por que a decisão tomada pelo sistema deve ser considerada uma decisão válida e bem fundamentada (Carbonera; Gonçalves; Souza, 2018).

Quanto mais sofisticada a máquina, pouco decifrável ela se torna para os humanos que dela dependem e um sistema de comportamento inteligente que decide sem explicar é como um juiz que sentencia em silêncio. E se a sentença não se abre à crítica, deixa de ser uma decisão jurídica. O fato de um modelo algorítmico ser interpretável não implica, contudo, que ele prescinda de mecanismos de explicabilidade (Nunes, 2023c).

A interpretabilidade se refere à facilidade de acesso ao funcionamento interno do modelo, mas a explicabilidade envolve a comunicação cognoscível e contextualizada das razões da decisão, especialmente para destinatários não técnicos. Entre as propriedades desejáveis em sistemas interpretáveis, destacam-se a transparência (vinculada à inteligibilidade

dos mecanismos internos do sistema) e a interpretabilidade *post-hoc*, associada à capacidade do sistema de fornecer informações relevantes sobre seus resultados a distintos perfis de usuários (Carbonera; Gonçalves; Souza, 2018).

Em contraste com os “modelos transparentes”, os chamados “modelos opacos” exigem um processo explicativo adicional (denominado “explicabilidade *post-hoc*”) para que suas decisões possam ser compreendidas (Nunes, 2023c). É uma tentativa de “traduzir” uma decisão que, por sua própria natureza, não foi pensada para ser comunicável, por isso, a adoção do termo “*post-hoc*” que em latim significa “depois disso”. Dentre os modelos considerados opacos, destacam-se as *Support Vector Machines* (SVM), o *Perceptron* multicamadas e as florestas aleatórias (*random forests*) (Nunes, 2023c).

No caso das florestas aleatórias, mencionadas no segundo capítulo, tem-se a atuação conjunta de múltiplas árvores de decisão, cujos resultados são agregados, o que dificulta atribuir uma justificativa clara a uma decisão específica. As SVMs operam em espaços matemáticos de alta dimensionalidade, baseando-se em funções de margem e hiperplanos que são opacas para usuários não especialistas e os perceptrons multicamadas, próprio nome indica a complexidade, utilizam o *deep learning* para seu funcionamento.

A opacidade algorítmica se manifesta de maneiras variadas. O primeiro nível se refere ao “sigilo corporativo ou estatal intencional”, imposto por empresas ou pelo Estado com o objetivo de preservar uma vantagem competitiva em relação a seus concorrentes (Nunes, 2023c). É um mistério intencional das corporações, muitas vezes sob a justificativa de proteger segredos comerciais ou de segurança nacional. O segundo nível de opacidade está relacionado ao “analfabetismo técnico”, uma vez que a escrita e a leitura de códigos algorítmicos exigem habilidades especializadas (Nunes, 2023c). Essa forma de opacidade não decorre de uma escolha estratégica de sigilo, mas da própria complexidade técnica dos sistemas de IA. Isso remete aos ensinamentos do primeiro capítulo, onde se esclareceu o funcionamento algorítmico que consistem em conjuntos de instruções que, quando seguidas, fazem o sistema operar de maneira automatizada.

A terceira se trata de uma opacidade ainda mais desafiadora do que as anteriores: a complexidade das operações internas dos sistemas de inteligência artificial (Nunes, 2023c). E coloca em questão a capacidade de monitorar, controlar e responsabilizar tais sistemas, tornando-se um

obstáculo significativo para a transparência e a confiança nas tecnologias emergentes. Nesses casos, torna-se imperativa a aplicação de um ou mais formas explicativas, a fim de que o modelo possa esclarecer suas próprias decisões (Nunes, 2023c).

É esse o objetivo das técnicas de explicabilidade *post-hoc* também chamadas de “explicabilidade pós-modelagem” que buscam transmitir informações acessíveis sobre o modo como um modelo já desenvolvido gera suas previsões para quaisquer entradas (Nunes, 2023c). O intuito não é modificar o modelo em si, mas proporcionar uma camada de entendimento adicional para aqueles que precisam validar ou questionar os processos decisórios. A explicabilidade *post-hoc* é aplicada a modelos cujo *design* não permite interpretação imediata, utilizando uma variedade de “métodos” e ferramentas para proporcionar explicações e entre esses recursos, destacam-se explicações textuais, visuais, exemplificativas, por simplificação e “análises” de relevância dos atributos (Arrieta, 2019).

As técnicas de explicabilidade *post-hoc* podem ser classificadas em duas categorias: as técnicas gerais ou agnósticas, que podem ser aplicadas a qualquer modelo de aprendizado de máquina, e as técnicas específicas, desenvolvidas para interpretar um modelo particular, o que impede sua extrapolação para outros modelos (Nunes, 2023c). Assim, a escolha entre uma abordagem agnóstica ou específica depende da necessidade de generalização da explicação que se busca alcançar. Dentre as técnicas agnósticas de explicabilidade, isto é, aquelas que podem ser aplicadas de forma geral a diferentes modelos, destacam-se as explicações textuais, que têm como objetivo representar o funcionamento do algoritmo por meio da linguagem escrita ou de símbolos (Nunes, 2023c). Com isso, torna-se viável entender os princípios lógicos que deram origem à decisões automatizadas.

Em segundo lugar, as explicações por simplificação figuram entre os métodos de explicabilidade mais utilizados, baseando-se na extração de regras que sintetizam o processo preditivo do modelo (Nunes, 2023c). Cria-se uma versão mais simples do modelo original como uma árvore de decisão que imita ou aproxima o comportamento do modelo complexo da floresta aleatória. Todavia, esse instrumento pode ser limítrofe, pois formas de aproximação e não são de revelação integral e tais explicações por vezes criam o risco de “transparência simulada.

A terceira técnica, a explicação visual, consiste na geração de gráficos, tabelas ou outros esquemas visuais pelo algoritmo, com o objetivo de tornar suas previsões mais compreensíveis ao usuário (Nunes, 2023c). Ao utilizar

visualizações, como a evidenciação dos atributos que influenciam decisões preditivas, por exemplo, o processo de aprendizagem se torna mais concreto e palpável. Outra técnica de explicabilidade aplicável a diversos modelos é a de relevância da informação, que busca evidenciar o peso atribuído a cada variável na formação de uma decisão algorítmica (Nunes, 2023c). Ao evidenciar o peso de cada variável na decisão automatizada, permite-se uma aproximação com a noção de justificação racional típica da dogmática jurídica. Por exemplo, no caso de um banco que nega um empréstimo a um cliente, essa técnica permitiria ao algoritmo indicar quais fatores foram determinantes para a recusa (Nunes, 2023c).

As explicações pontuais têm como objetivo esclarecer segmentos específicos (denominados “locais”) do modelo, destacando partes do seu funcionamento que são particularmente relevantes para determinados resultados (Nunes, 2023). A explicação pontual se concentra em um resultado isolado, investigando quais características (variáveis ou entradas) foram mais determinantes para aquela saída específica. A sexta e última técnica de explicabilidade é a utilização de exemplos (Nunes, 2023c). Em vez de tentar explicar tecnicamente o funcionamento interno do modelo, essa abordagem mostra exemplos concretos que serviram de base ou referência para a decisão tomada.

Já, a seleção das técnicas específicas varia conforme o tipo de modelo utilizado, a natureza do problema enfrentado e o perfil do público a que se destinam os resultados (Marques; et. al, s.d). Essa tríade evidencia que a explicabilidade é uma prática relacional, com seis opções que se destacam. Começando pela visualização de coeficientes que é aplicável em modelos lineares, como a regressão linear ou a regressão logística, pode-se visualizar diretamente os coeficientes atribuídos a cada característica, o que permite destacar a influência de cada uma sobre a variável dependente (Marques; et. al, s.d).

Com os gráficos de barras que representam a importância das características, é possível criar diagramas que ilustram a relevância relativa de cada variável no processo decisório, especialmente em modelos como árvores de decisão e outros baseados em regras (Marques et al., s.d.). Ao revelar quais variáveis influenciam decisivamente as saídas do modelo, essas representações permitem aferir se a decisão está sendo orientada por critérios coerentes com os valores e normas aplicáveis em determinado contexto.

Por seu lado, os gráficos de dependência parcial ilustram como a variável de saída se relaciona com uma variável de entrada específica, mantendo as demais variáveis constantes (Marques et al., s.d.). Enquanto os primeiros indicam, de forma global, quais variáveis possuem maior peso na construção das decisões do modelo, os segundos permitem observar, de modo mais detalhado e contextual, a direção e a intensidade da relação entre variáveis e previsões.

As matrizes de correlação, em modelos interpretativos como as regressões lineares, pode auxiliar na identificação de relações lineares entre as variáveis explicativas e a variável de saída (Marques et al., s.d.). Permite observar quais variáveis estão linearmente associadas entre si, isto é, se o aumento de uma variável tende a acompanhar o aumento (ou a diminuição) de outra. A “análise” de resíduos, nos modelos de regressão, contribui para a avaliação das suposições do modelo, como a linearidade e a homocedasticidade (Marques et al., s.d.). Se as diferenças entre os valores observados estiverem distribuídas de forma aleatória e sem padrão, isso indica que a relação é realmente linear. A regressão linear pode ser empregada como uma ferramenta estatística para sintetizar dados e informações de forma objetiva (Chein, 2019). A homocedasticidade significa que a variância dos erros deve ser constante para todos os valores das variáveis independentes.

Em modelos baseados em regras de explicação, as decisões podem ser justificadas por meio de enunciados simples do tipo “se... então...”, o que facilita a compreensão do funcionamento do modelo (Marques et al., s.d.). Esse formato lógico tende a clarificar a forma de raciocínio ao tornar explícita a relação entre premissas e conclusões. Ao aplicar isso no campo do direito processual civil, aplica-se a seguinte estrutura: “se a parte não apresentar contestação no prazo legal, então será decretada a revelia”.

Em síntese, as categorias de explicabilidade representam distintos modos pelos quais se busca compreender, justificar ou tornar inteligível determinado fenômeno, conceito ou decisão. A seguir, apresenta-se um breve resumo dessas categorias, conforme anteriormente exposto, com o propósito de sistematizar os principais aspectos abordados.

Figura 04: Distinção entre técnicas gerais/agnósticas e específicas de explicabilidade.



Fonte: Autora (2024)

Nessa perspectiva, é inviável abordar a questão da opacidade e dos vieses algorítmicos sem considerar a transparência e, sobretudo, a necessidade, no âmbito jurídico, de incorporar princípios e normas destinados a regular a atuação algorítmica nas diversas áreas em que se faz presente (Nunes, 2023c). Trata-se de assegurar que os sistemas de comportamentos inteligentes, ao ingressarem em processos decisórios que afetem direitos fundamentais, estejam submetidos a exigências de justificação, explicabilidade e controle público. Em primeiro lugar, porque o aumento do conhecimento público sobre os processos algorítmicos favorece um debate mais informado e crítico acerca da implementação e evolução das novas tecnologias (Nunes, 2023c).

Em segundo lugar, porque contribui para o aprimoramento das decisões geradas por sistemas de IA, ao permitir a identificação de *outputs* discriminatórios e a mitigação de vieses presentes nos modelos algorítmicos (Nunes, 2023c). A preocupação não é apenas com o conteúdo da decisão, mas com os processos que envolvem sua fundamentação. O desenvolvimento do “direito à explicação” e o aprimoramento das técnicas e premissas da explicabilidade podem trazer benefícios tanto para os indivíduos quanto para a sociedade como um todo, uma vez que promovem a criação de sistemas mais transparentes e interpretáveis, aumentando a confiabilidade e a legitimidade das decisões geradas por máquinas (Nunes, 2023c).

Nessa maré, o Projeto de Lei 2.338/2023, aprovado no último mês do ano de 2024, trouxe em seu escopo princípios voltados à governança ética da inteligência artificial, com especial destaque para a transparência e a explicabilidade dos sistemas automatizados. Como já abordado no primeiro capítulo desta dissertação, no qual se discutiram os principais marcos regulatórios sobre o tema, a regulação da IA tem buscado responder aos desafios da opacidade algorítmica por meio da consolidação de princípios jurídicos que assegurem maior controle e responsabilidade sobre os sistemas utilizados.

Nesse contexto, o Artigo 3º do projeto de lei destaca a transparência e a explicabilidade como princípios fundamentais, levando em consideração o segredo comercial e a participação de todos os agentes envolvidos na cadeia de valor, além de garantir a diligência devida e a auditabilidade durante todo o ciclo de vida dos sistemas, conforme o estado da arte do desenvolvimento tecnológico e o grau de risco associado. O projeto também ressalta a confiabilidade e robustez dos sistemas de IA, com o objetivo de assegurar a proteção dos direitos fundamentais, como o devido processo legal, a contestabilidade e o contraditório, enfatizando a prestação de contas, a responsabilização e a reparação integral de danos. Por fim, o mesmo artigo reforça a importância da prevenção, precaução e mitigação de riscos e danos, exigindo que os métodos empregados sejam proporcionais às finalidades legítimas dos sistemas, com foco na proteção dos direitos fundamentais individuais, sociais, coletivos e econômicos.

Identificaram-se os principais objetivos que devem ser perseguidos para a efetivação da explicabilidade (Nunes, 2023c). Trata-se de critérios que não apenas orientam o desenvolvimento técnico dos modelos, mas que também possuem implicações diretas para a legitimidade jurídica e epistêmica de suas decisões. Dentre eles, destacam-se: a) causalidade, que busca evidenciar a correlação causal entre as variáveis envolvidas no processo decisório; b) transferibilidade, relacionada à capacidade de um mesmo modelo de IA ser aplicado em diferentes sistemas, bem como às limitações dessa transposição; c) informatividade, que consiste em fornecer informações detalhadas sobre o problema enfrentado pela máquina, considerando que ele nem sempre corresponde ao problema percebido pelo usuário; d) confiança, pela qual o modelo deve apresentar elementos que permitam ao usuário avaliar a robustez e a estabilidade de seu funcionamento e de suas decisões; e) equidade, que possibilita uma “análise” ética e justa das decisões tomadas pelo sistema; f) acessibilidade, referente à possibilidade de os usuários finais participarem do processo

de aprimoramento e desenvolvimento do modelo; g) interatividade, que permite ao modelo dialogar com o usuário final, promovendo uma compreensão mais aprofundada; e h) conscientização da privacidade, que assegura ao usuário o entendimento sobre eventuais violações de seus dados pessoais pelo algoritmo (Nunes, 2023c).

Nesse sentido, a explicabilidade não pode ser reduzida a um atributo técnico, pois se projeta como um verdadeiro pressuposto processual e garantístico, assegurando que as decisões mediadas por IA permaneçam comprometidas com os valores do Estado Democrático de Direito e com a integridade da prática jurídica. A sedutora promessa de neutralidade do julgador e de celeridade na prolação das decisões se apresentaria de maneira inigualável, diante da exclusão da participação humana no julgamento e da instantaneidade com que a máquina processa as informações (Momolli, 2018). A solução aparentemente eficiente, ignora as complexidades do papel humano no processo decisório. Nesse sentido, ressurgiria a pretensão da filosofia racionalista de eliminar o subjetivismo inerente ao julgador (Momolli, 2018).

Soma-se a isso o endeusamento da técnica enquanto forma de ver o mundo, o que, em verdade, apenas recicla antigas concepções filosóficas: mudam-se os rótulos (“algoritmos” soa mais sedutor que “empirismo”), mas o conteúdo permanece essencialmente inalterado (Streck; Bernsts; Gomes, 2021). É uma tendência de vestir antigas concepções filosóficas com trajes modernos. No processualismo civil isso se traduz na progressiva redução do processo a uma engrenagem tecnocrática, orientada por formalismos procedimentais e por uma lógica maquinal. Enquanto a computação tende a supervalorizar a capacidade de prever o que ocorrerá, acaba por negligenciar a aptidão de antecipar o que pode ocorrer (Carbonera; Gonçalves; Souza, 2018).

Em textos anteriores, já se apontou o paradoxo: caso a extrema tecnologização obtenha êxito, dará errado (Streck; Bernsts; Gomes, 2021). Caminha-se para a beira do naufrágio da justiça que se perde enquanto fenômeno humano, dialógico e aberto à riqueza da hermenêutica. Assim como ocorre nas tentativas de objetivar ou matematizar o cérebro e as emoções por meio de eletrodos e dispositivos análogos, se funcionarem aniquilam precisamente aquilo que pretendiam compreender, minando o espaço da filosofia (Streck; Bernsts; Gomes, 2021).

Acredita-se que, a partir do mapeamento cerebral e da possibilidade de codificar tal processo, tornar-se-á possível delinear com maior precisão

o conceito de mente e, por conseguinte, o de ser humano (Streck; Jung, 2024). Logo, ao pretender definir a essência do ser humano, assume-se a ideia de que a mente (consciência, subjetividade e experiência) pode ser reduzida a um simples conjunto estrutural de dados, passível de controle e manipulação. Há uma redução da humanidade a um ente mecânico, desconsiderando sua essência plural, histórica e moldada por experiências complexas. Com isso, estados de consciência e suas implicações práticas poderiam ser previstos matematicamente e, até mesmo, reproduzidos artificialmente (Streck; Jung, 2024).

Uma vez alcançado esse estágio, hoje considerado iminente, os debates de natureza mística ou secular sobre os diversos aspectos controversos da existência humana, tidos como meros epifenômenos, tenderiam a perder relevância ou a se encerrar (Streck; Jung, 2024). Este movimento de superação de discussões mais filosóficas ou metafísicas pode ser compreendido à luz das contribuições de pensadores como René Descartes, que, em sua obra *Discurso do Método*, já defendia a ideia de que a natureza humana é completamente independente do corpo. Para Descartes, a alma não está sujeita à morte do corpo e, portanto, é imortal.

Argumentava-se que a alma racional não poderia ser derivada do poder da matéria, mas deveria ser criada de forma distinta, íntima e inseparável do corpo humano. Esta visão implica que a alma não é uma entidade meramente material, mas um princípio que confere à pessoa sentimentos, apetites e a capacidade de compor o que verdadeiramente é um “homem”. Dessa forma, Descartes via a natureza humana como algo que transcende o material, cuja compreensão exige algo além do que os olhos podem ver ou a ciência pode explicar plenamente.

Neste contexto, a citação do espiritualista Alan Kardec e a conceituação da palavra “lei” na Bíblia Sagrada se justificam na presente dissertação não com o intuito de incidir um viés religioso, mas para ressaltar a necessidade de se reconhecer o ser humano em sua plenitude e complexidade. Brillantemente, Ovídio Batista denomina o Direito como a “ciência do espírito”, uma expressão que resgata a dimensão mais profunda e imaterial da existência humana, algo que vai além das definições reducionistas e mecanicistas frequentemente associadas à prática jurídica. À vista disso é imprescindível posicionamentos não em oposição, mas de forma crítica e reflexiva diante da crescente “colonização do universo jurídico” pelos avanços tecnológicos (Streck; Bernsts; Gomes, 2021).

Diante do avanço inevitável da técnica, enseja-se a investigação nas tomadas de decisões do poder judiciário (Momolli, 2018). A técnica deve estar a serviço da justiça e não o contrário. Atualmente, em um contexto profundamente marcado por táticas e estratégias, torna-se insuficiente afirmar que a técnica é boa ou má a depender do uso que dela fazemos (Tybusch, 2025). Não se trata de afirmar a neutralidade da técnica ou sua submissão à vontade humana, mas de reconhecer seu papel constitutivo no mundo e nas relações que o atravessam. Rompe-se com uma concepção instrumental clássica.

Nesse contexto, a técnica deixa de ser apenas um instrumento sob controle e se transforma no próprio meio em que as pessoas estão imersas, de tal forma que já não há o questionamento sobre como a utilizá-la ou o que fazer com ela, mas se passa a indagar sobre o que a técnica pode fazer com a sociedade (Tybusch, 2025). A interrogação desloca o foco do sujeito que domina para o sujeito que é por ela condicionado. Essa inversão é central para (re)pensar as implicações éticas, políticas e jurídicas da tecnicidade em sociedades marcadas pela automação e pela racionalidade operacional.

Perceber o funcionamento da técnica demanda um exercício contínuo de atenção e discernimento, de modo que, guiados pela prudência, tenha-se capacidade de considerar as particularidades de cada situação e optar pelas aplicações mais apropriadas às circunstâncias (Tybusch, 2025). Em vez de se adotar uma confiança cega na técnica ou, em sentido oposto, sua simples rejeição, é necessário recorrer à *phronesis* uma forma de sabedoria prática que se orienta pela prudência e pela consideração das singularidades de cada situação. Nesse sentido, a responsabilidade pelo uso da técnica é inseparável de uma atitude reflexiva, crítica e atenta ao humano, capaz de resistir à tentação de soluções fáceis, padronizadas ou previamente formatadas.

As decisões, portanto, devem ser construídas no caminho, e a escolha mais legítima será sempre aquela mais humanizada (Tybusch, 2025). Não se trata de aplicar aprendizados de máquinas ou fórmulas universais, mas de responder com responsabilidade ao outro, considerando sua condição concreta, sua história e sua vulnerabilidade. Por isso, indaga-se se algoritmos rimam com epistemologia? Como a IA lida com aquilo que é condição para explicarmos o que compreendemos? Ou também já matamos a epistemologia? (Streck; Bernsts; Gomes, 2021).

Ainda que se admita que máquinas e juízes possam decidir juntos, a máquina nunca terá a capacidade de compreender o fenômeno complexo

que se desenvolve a partir de um processo, pois interpretar é, por si, ato hermenêutico complexo, que depende da atividade humana, vinculada à capacidade de compreender a historicidade e desvelar a tradição (Streck; Bernsts; Gomes, 2021). A preocupação com a explicabilidade *post-hoc* expressa, assim, um esforço de manter a responsabilidade humana no processo decisório.

Apesar de se encaixar no crivo técnico, trata-se de uma exigência crítica mínima sobre um modo de resgatar a dimensão ética e narrativa do direito diante do risco de automatização desumanizante. A crítica de Lenio Streck ao decisionismo e ao tecnicismo é central nesse ponto, pois evidencia que decisões judiciais não podem ser resultado de um juízo técnico ou discricionário, mas exigem uma tomada de posição comprometida com a Constituição e com a dignidade da pessoa humana. Aliás, a formação de um precedente constitui (ou ao menos deveria constituir) uma obra humana e coletiva, resultante do julgamento de diversas pessoas e destinada à aplicação por outras em momentos subsequentes (Streck, 2025).

A jurisdição humana, nesse contexto, não é uma opção entre outras, mas uma exigência do próprio desenho constitucional do processo. Ademais, carece da aptidão para reconhecer o argumento de princípio que estabelece a conexão entre precedentes (Streck, 2025). A tentativa de delegar essa função a sistemas automatizados, sem o devido protagonismo humano, compromete a legitimidade do processo e empobrece a prática jurídica. Um robô pode, de fato, incorrer em menos erros que um ser humano ou demonstrar maior consistência e menor subjetivismo; contudo, permanece incapaz de reproduzir sentimentos genuinamente humanos, como a piedade (Streck, 2025).

O direito fundamental à jurisdição humana, pautada por critérios e obediente aos pressupostos processuais estabelecidos em dispositivos como o artigo 93, inciso IX, da Constituição Federal, o artigo 926 do Código de Processo Civil, e os artigos 489 do CPC e 315 do Código de Processo Penal (Streck, 2025). Esses dispositivos não podem ser vistos como meras formalidades, mas como garantias constitucionais, que se opõem tanto ao decisionismo quanto à automatização (a)crítica promovida pela aplicação cega/opacas de tecnologias.

Desse modo, sair da Matrix não significa propriamente abandoná-la ou libertar-se dela, mas sim “tomar consciência”, reconhecer que estamos imersos nela (Espindola, 2022). Passa-se a indagar sobre o que a técnica pode fazer com a sociedade. Nesse sentido, a metáfora da *Matrix*

ganha força quando situada no horizonte do *perguntar*, tal como proposto por Heidegger porque perguntar é abrir o ser àquilo que se esconde, é desvelar o encobrimento da existência cotidiana. A técnica obscurece a essência das coisas, reduzindo o ser ao que pode ser mensurado ou melhor: automatizado. Para tanto, é necessário decodificá-la e desvelá-la, para então reiniciá-la (refundá-la) sob um novo paradigma (Espindola, 2022).

Nesse horizonte caboteiro, a libertação ou mesmo uma aventura *nirvânica* (entendida como chave mestra para o despertar das construções ilusórias e para a superação do sofrimento e da ignorância) não consiste em rejeitar a técnica, mas em desativar sua dominação invisível por meio da consciência crítica. A título ilustrativo, e com base na metáfora proposta, destaca-se a cena do filme em que se afirma: “*libere sua mente, para pular de um prédio para outro*” (Espindola, 2022).

A assertiva sugere que, mesmo estando “dentro” da Matrix, é possível “estar fora” dela ao nos apropriarmos das habilidades humanas, da artesanaria do processo, expressão que cunhamos para designar a observância das garantias constitucionais na interpretação e aplicação do Direito e atuarmos como outsiders, ainda que inseridos em uma Matrix previamente decodificada (Espindola, 2022). Trata-se, pois, de uma convocação a resistir ao automatismo das decisões e à normatividade fria para que não se crie uma sociedade de juristas mortos, como adverte Lênio Streck, reafirmando o papel do intérprete comprometido com os fundamentos constitucionais e com os direitos fundamentais.

Buscando manter a coerência com a metáfora proposta, vale lembrar que, em *Matrix: Reloaded*, quando Neo encontra o Arquiteto (criador da Matrix), descobre que ele próprio é uma peça intencional do sistema, concebido para evitar seu colapso (Espindola, 2022). A figura do Arquiteto, como aquele que projeta a ilusão do livre-arbítrio dentro de um sistema fechado, representa uma crítica pungente à racionalidade tecnocrática que permeia o Direito atual. Neo, ao descobrir que é parte programada do sistema, simboliza a inquietante ideia de que até mesmo os agentes da ruptura podem ser engolidos por uma lógica de auto reprodução estrutural. Diante disso, é confrontado com uma escolha: retornar à fonte e reiniciar a Matrix ou destruí-la, desconectando e, consequentemente, condenando todos os que nela estão (Espindola, 2022).

O dilema da escolha, nesse contexto, simboliza a lógica binária, a linearidade e os padrões próprios da linguagem de programação (Espindola, 2022). É o ponto de tensão entre a velha e saudosa alma jurídica ombreada

pelas bibliotecas empoeiradas e das inquietações humanas mais profundas, com os circuitos frios da modernidade tecnocrática. Lembra-se que para Parmênides, os mortais incorreram em erro ao conceber a distinção entre o ser e o não-ser, e ao erigir em duas formas supremas “luz” e “noite” como antagônicas. O filósofo refutou a ideia de que a realidade se construa a partir de oposições porque a essência da realidade é desvinculada de qualquer dualismo.

Já, Heráclito sustentava que o verdadeiro conhecimento só poderia ser atingido por meio do autoconhecimento e da compreensão profunda da natureza do cosmos. Para ele, ao compreender a linguagem (sua dinâmica e seus significados) seria possível alcançar uma compreensão mais profunda da própria realidade. Se o Arquiteto representa o criador da Matrix (uma metáfora para o sistema jurídico) e Neo encarna o intérprete “irrevogavelmente humano”, então, ainda que o processo tenha transformado sua consciência, como se anuncia no diálogo com o Arquiteto, não é a linguagem binária que determina ou comanda o intérprete, mesmo que possa auxiliá-lo na tarefa de reduzir as complexidades do mundo jurídico (Espindola, 2022)

Embora o Arquiteto seja o criador da Matrix, é o intérprete que programa o Arquiteto e sua linguagem binária (Espindola, 2022). A linguagem do Arquiteto só comunica algo porque é interpretada. No contexto da jurisdição e, por consequência, do acesso à justiça, a indagação persiste, ainda que com nova roupagem para quem programa o programador? (Espindola, 2022). Afinal, o “programador” (seja ele o engenheiro de software, o legislador que autoriza o uso de inteligência artificial ou a empresa que fornece a solução) carrega consigo valores, escolhas e visões de mundo que se infiltram na lógica do sistema.

O intérprete não é um elemento externo ao sistema, mas sim aquele que, como Neo, ao tomar consciência de sua própria inserção nele, pode tensionar os limites do possível e reconfigurar o próprio sentido da justiça. Nesse contexto, vale lembrar que foi objetificada, desde Platão, o próprio conceito de existência. O que é o Estado? O que é a justiça? Isso são entes. E a conversão algorítmica só se concretiza quando se adere ao paradigma sujeito-objeto e à rigidez metodológica, pois ela envolve antecipações e o exercício de discricionariedade (Momolli, 2018). Abre-se um parêntese para dizer que a vida e o ser humano são reduzidas a um objeto mercantil, através das lentes de um tecnoliberal.

Segundo Eduardo Grüner, a pergunta heideggeriana sobre a técnica revela “uma lógica cuja finalidade é substituir a Verdade do Ser por um saber mecanicista, que transforma o mundo em uma imagem eficiente, porém desprovida de fundamento e valor profundo” (Mavrakis, 2021). A vertente positivista do direito desassocia o indivíduo de sua realidade concreta (condições materiais, inserção social e dificuldades de acesso a direitos fundamentais), reduzindo-o à figura abstrata de um sujeito processual delimitado apenas pela ação judicial que propõe ou sofre.

O Direito deve considerar a condição ontológica do sujeito diante do emaranhado de sentidos, visto que o plano ôntico diz respeito àquilo que já está dado no mundo, como elementos que “mobíliam” o nosso entorno. E o plano ontológico remete ao mais essencial que são as ações em si. Assim, ao pensar o Direito, é imprescindível não apenas lidar com os elementos ônticos (as normas, os procedimentos, os códigos), mas também considerar sua dimensão ontológica, ou seja, as condições de possibilidade do sentido jurídico diante da existência humana concreta.

Há autores que descurem o fato de que a explicabilidade não se configura como uma noção absoluta, mas uma construção contextual (Carbonera; Gonçalves; Souza, 2018). Vê-se que debates longínquos, outrora circunscritos à filosofia, aos estudos da mente (psicologia e a neurociência), à cibernética e à linguagem, retornam ao centro das discussões contemporâneas com o avanço crescente da inteligência artificial. O quadro delineado no primeiro capítulo evidencia essa reatualização temática, demonstrando como questões clássicas vêm sendo revisitadas sob novas perspectivas tecnológicas e epistemológicas. Existe distinção entre o símbolo (a palavra em si) e seu significado (sentido pragmático em contexto específico) (Carbonera; Gonçalves; Souza, 2018).

Não obstante, as divisões de competências, interesses, formações e práticas profissionais, tanto no campo científico quanto na esfera do desenvolvimento tecnológico, reiteram o hiato histórico que separou o estudo do uso dos sistemas linguísticos, sejam eles naturais ou artificiais, da investigação de sua estrutura ou lógica subjacente (Carbonera; Gonçalves; Souza, 2018). Esse resgate não ocorre de forma dissociada, mas encontra ressonância direta com a explicabilidade *post-hoc*, isto é, aquela que busca oferecer justificativas compreensíveis e retrospectivas para as decisões produzidas por sistemas algorítmicos opacos. A necessidade de tornar inteligíveis os processos decisórios das máquinas reativa discussões sobre

consciência, intencionalidade e linguagem, revelando uma continuidade sofisticada entre o antigo e o emergente.

O aspecto mais recorrentemente salientado reside no fato de que a confiança nas decisões e avaliações geradas por tais sistemas depende, assim como ocorre entre os próprios seres humanos, da capacidade de o usuário (seja beneficiário ou de algum modo impactado pelo resultado da inferência automática) se convencer de sua legitimidade e fundamentação (Carbonera; Gonçalves; Souza, 2018). Se explicar o sentido da vida já se mostra uma tarefa árdua e encontrou rotas de fuga convenientes, como Deus ou o mundo metafísico das ideias.

Tais explicações, embora consoladoras, muitas vezes servem mais como subterfúgios do que como respostas efetivas à complexidade da existência. De modo análogo, a explicabilidade *post-hoc* também pode ser uma saída técnica para a tomada de decisão judicial automatizada tendo em vista que possibilita alguma forma de reconstrução racional das inferências realizadas, favorecendo o contraditório e a legitimidade decisória.

Entretanto, a crescente presença de robôs e ferramentas automatizadas embora apresentada como resposta à morosidade e ao excesso de processos não tem revertido a sobrecarga judiciária. Conforme dados citados em linhas anteriores, o número de demandas continua a crescer, o que sugere que o problema é menos técnico e mais humano, institucional e estrutural. A comunicação, por si só, já é algo improvável, logo, o papel da IA nas narrativas se revela ainda mais desafiador. Por isso, refundar a jurisdição não implica destruí-la (como propõe a “*destruktion*” heideggeriana), mas sim desconstruí-la, à maneira de Derrida, para então reconstruí-la a partir de seus fundamentos, especialmente considerando o paradigma da linguagem no qual a filosofia já se inscreve e que o Direito, ainda de forma tardia, busca acompanhar (Espindola; Silva, 2018).

Especialmente a partir do pensamento de Owen Fiss para quem, ainda que próximo às perspectivas procedimentalistas e substancialistas, é possível conceber uma atuação transformadora do Judiciário por meio de uma nova forma de adjudicação: a reforma estrutural, Arenhart e Vitorelli Lima vêm propondo uma abordagem voltada à efetividade da tutela jurisdicional coletiva no Brasil, denominada processo estrutural (Momolli, 2020). Visa-se reverter a lógica de uma justiça desprovida de substância. Afinal, o direito está inserido em um contexto social complexo, no qual a mera existência de regulamentação jurídicas se revela insuficiente se não acompanhada de outras medidas complementares, que vão desde

uma adequada organização institucional até o respeito dos cidadãos, entre diversos outros aspectos (Streck, 2015).

Dificulta-se a previsão do desfecho final da trajetória com a inteligência artificial porque não se trata apenas de uma narrativa sobre máquinas, mas também de escolhas, de liberdade e da possibilidade de moldar o próprio destino (Lee, 2019). A inteligência artificial pode acabar operando como um deslocamento ilusório: uma mudança de meio que preserva a mesma lógica saturada, sem transformá-la de fato. Além disso, enquanto o ser humano permanecer aprisionado em conceitos inibidores da alma corre o risco de não perceber que a vida ultrapassa aquilo que é puramente visível, mensurável ou computável. Despertar (da matrix) exige mais do que técnica: exige consciência. Nesse processo, que se busque olhar profundamente para dentro e para o outro, em busca dos valores e da sabedoria capazes de orientar o caminho (Lee, 2019).

5 CONCLUSÃO

“Não conheço o futuro. Não vim aqui para te dizer como isso acabará, vim para dizer como começará.”

– Neo, em *Matrix* (Wachowski; Wachowski, 1999)

A investigação proposta lança luz sobre a complexa interseção entre inteligência artificial e os fundamentos éticos, ontológicos e institucionais do Direito, inspirando-se na metáfora cinematográfica de *Matrix* para descortinar os mecanismos invisíveis de controle que perpassam a sociedade digital contemporânea. Parte-se do pressuposto de que compreender o presente exige não apenas o domínio técnico das novas tecnologias, mas sobretudo a recuperação dos marcos históricos da ciência.

À vista disso, iniciou-se o percurso teórico com base em autores como Allen Newell, Herbert Simon, Norbert Wiener, Ashby, Castells, Virilio e outros, demonstrando que a ciência é um fenômeno histórico que avança pela compreensão das experiências pretéritas. A denominada sociedade em rede, que caracteriza o contexto contemporâneo, assenta-se sobre estruturas organizacionais pretéritas, razão pela qual não se a qualifica simplesmente como sociedade da informação ou da comunicação, fenômenos estes que sempre acompanharam a trajetória humana.

O traço verdadeiramente inovador reside na emergência das redes digitais e nas novas configurações jurídicas que delas decorrem. Surge um cenário de desmaterialização das referências tradicionais de espaço e representação, marcando uma transição da arquitetura para uma “pós-arquitetura”, conforme aponta Virilio. Os antigos limites físicos, como portões, são substituídos por barreiras digitais, reguladas por bancos de dados e redes de informação. Desse modo, desmistifica-se a visão simplista e instrumental do termo “algoritmo”, demonstrando como esse conceito fundamenta tanto práticas computacionais quanto padrões de racionalidade humana. Assim, estabelece-se a expressão adequada para se referir a sistemas de comportamentos inteligentes, sendo o teste de Turing um marco seminal na definição operacional da inteligência em máquinas.

Apesar de ganhar maior visibilidade nos últimos anos, o panorama histórico das 27 principais questões intelectuais no campo da Inteligência Artificial inicia entre os séculos XVII e XX. A história das ciências cognitivas

e da inteligência artificial foi marcada por uma série de dicotomias conceituais que moldaram sua evolução. A substituição da teleologia pelo mecanicismo, consolidada com o surgimento da cibernética, abriu caminho para uma trajetória marcada por separações estruturantes: entre razão e emoção, biologia natural e vitalismo, filosofia e ciência da mente, lógica e psicologia. Com o advento da computação e da IA simbólica, novas clivagens emergiram como símbolos versus números, heurísticas versus algoritmos, desempenho versus aprendizagem, sintaxe versus semântica, esta última responsável por dissociar o estudo da linguagem dos sistemas de IA. Delimita-se, assim, territórios disciplinares.

Adentrando o núcleo técnico da IA, o texto explora a definição de agente inteligente, combinando arquitetura (hardware + sensores e atuadores) e programa (software). São abordadas as principais vertentes do aprendizado de máquina: aprendizado supervisionado, não supervisionado e por reforço. Também foi apresentado o aprendizado profundo (*deep learning*), que utiliza redes neurais artificiais inspiradas na estrutura do cérebro humano para extrair padrões e realizar inferências complexas. A arquitetura dessas redes neurais (com camadas de entrada, ocultas e de saída) revela um processamento cada vez mais sofisticado.

Avança-se para distinção entre inteligência artificial fraca (ANI), forte (AGI) e superinteligência (ASI), apontando para o desafio existencial representado pela possível emergência de inteligências artificiais que ultrapassem exponencialmente a cognição humana. A IA fraca já está presente no cotidiano (como em assistentes virtuais e jogos), enquanto a IA forte e a ASI permanecem em estágio hipotético, mas com potencial disruptivo. Desse modo, a urgência de regulamentações éticas é posta como contraponto à tendência de invisibilização dos riscos envolvidos, especialmente no que tange à autonomia humana. Como a arquitetura da cidade moderna cede lugar a uma urbanidade informacional, o Direito também se transforma, exigindo dos juristas um compromisso ético com a compreensão dos sistemas inteligentes que ora permeiam a decisão judicial.

Desse modo, foram trazidas à baila as principais abordagens regulatórias criadas, com especial ênfase nas legislações da União Europeia, do Brasil e dos Estados Unidos. Na União Europeia, destaca-se o *AI Act* (Regulamento 2024/1689), aprovado em maio de 2024, como o primeiro instrumento jurídico vinculante do mundo a classificar sistemas de IA conforme níveis de risco (inaceitável, alto, limitado e mínimo) estabelecendo obrigações proporcionais à sua periculosidade. O arcabouço é sustentado

por valores fundantes do bloco europeu, como a dignidade humana, o Estado de Direito e a proteção ambiental, promovendo um modelo de IA centrado no ser humano e comprometido com a transparência, a auditabilidade e a explicabilidade dos sistemas.

No Brasil, o Projeto de Lei nº 2.338/2023, aprovado pelo Senado em dezembro de 2024, apresenta diretrizes gerais para a governança responsável da IA, incorporando princípios como a centralidade da pessoa humana, a proteção de dados, a não discriminação, a inovação segura e a responsabilidade solidária entre desenvolvedores, distribuidores e aplicadores. A proposta brasileira inova ao adotar um modelo mais horizontal de responsabilização e ao prever mecanismos como o sandbox regulatório e a categorização de riscos em dois níveis: risco excessivo e alto risco. Foi inspirada diretamente no *AI Act*, assim como a Lei Geral de Proteção de Dados (LGPD) se inspirou no Regulamento Geral de Proteção de Dados da UE (RGPD)

Nos Estados Unidos, por outro lado, observa-se uma abordagem regulatória fragmentada e não centralizada. Iniciativas como o *Blueprint for an AI Bill of Rights* (2022) e a ordem executiva de Joe Biden (2023, posteriormente revogada em 2025) demonstram esforços pontuais, mas ainda insuficientes para a constituição de um marco federal abrangente. A regulação norte-americana permanece fortemente influenciada por interesses das big techs, como Meta, Google e Microsoft, e por uma tradição de menor intervenção estatal, o que tem gerado preocupações relativas à segurança, à transparência e à equidade. No plano internacional, destaca-se a Declaração de Bletchley (2023), da qual o Brasil é signatário, documento que reconhece oportunidades e riscos sistêmicos da IA e propõe mecanismos de cooperação internacional. A declaração aponta para a urgência de coordenação transnacional na mitigação dos impactos da IA sobre a segurança, a soberania e os direitos humanos.

Ao seguir o modelo europeu, o Brasil sinaliza sua disposição em alinhar-se a padrões internacionais de proteção e inovação tecnológica. No entanto, o êxito dessas regulamentações dependerá de sua capacidade de resistir às pressões mercadológicas e de se afirmar como instrumentos de transformação estrutural, e não meros dispositivos legitimadores de um novo regime de poder algorítmico, como de maneira distinta ocorreu no processualismo civil brasileiro.

Parte-se da compreensão de que o processo judicial, enquanto manifestação da racionalidade ocidental, carrega consigo marcas profundas

da trajetória do Estado moderno (do liberalismo ao Estado Democrático de Direito) e que sua lógica formalista ainda reflete valores e estruturas históricas insuficientes diante das exigências de um mundo tecnológico. O positivismo jurídico, influenciado pelas ciências exatas e pelo ideal de objetividade, construiu um processo centrado em etapas e fórmulas, esvaziando o valor do caso concreto e afastando a justiça de sua dimensão hermenêutica. Como consequência, observa-se um Judiciário moroso que diante de suas próprias limitações, busca com a virada tecnológica uma resposta à crise de efetividade e legitimidade que enfrenta. Ao passo que essa é marcada por três fases principais: virtualização, automação e transformação do modo de agir.

O Judiciário brasileiro ingressa assim em uma nova era, com a introdução de plataformas como o Sinapses, o InovaPJe, o PDPJ-Br (instituída pela Resolução nº 335/2020 do CNJ), o Juízo 100% Digital, o Balcão Virtual, e o uso crescente de modelos de inteligência artificial treinados por sistemas supervisionados, não supervisionados e por reforço. A inteligência artificial é empregada em tarefas como categorização de peças processuais, previsão de desfechos judiciais e sumarização textual, utilizando modelos como redes neurais, árvores de decisão, florestas aleatórias e máquinas de vetores de suporte. Destaca-se a profusão de robôs e sistemas inteligentes espalhados por praticamente todos os tribunais brasileiros, com destaque para o Sinapses, e robôs como Victor, Rafa, Vitória, MarIA (STF), Athos (STJ), Assis (TJRJ), Oxóssi, Sofia, Jerimum, Haia, Poti, Arandu, entre os 147 presentes.

Ainda, existem parcerias com universidades e empresas privadas como Jusbrasil, Microsoft, Google, Softplan e AWS que têm gerado ecossistemas inovadores, com destaque para projetos de IA generativa e modelos de linguagem (LLMs), como o Jus IA, capaz de produzir petições e pareceres a partir da descrição dos fatos. O campo do Visual Law e do Legal Design proporciona abordagens que unem direito, *design* e linguagem simplificada para tornar os documentos jurídicos mais acessíveis, claros e eficazes. A partir de recursos como infográficos, linhas do tempo, vídeos e animações, o direito torna-se hipermodal, alinhando-se ao modo de leitura contemporâneo e à sobrecarga do Judiciário.

Outro eixo central são os meios alternativos de resolução de conflitos digitais, as chamadas “ODRs” (*Online Dispute Resolution*). A Resolução nº 354/2020 e a Resolução nº 358/2020 do CNJ instituíram normas para audiências de mediação online e sistemas informatizados de

resolução consensual (SIREC), estimulando a criação de plataformas como Semprocesso, Projuris, Mediar360, Consumidor.gov, MOL, Resolvvi, entre outras. As ODRs são vistas como uma “quarta parte” da solução de conflitos, ao lado do autor, do réu e do mediador humano, trazendo algoritmos e IA para prever desfechos e auxiliar na autocomposição. O IRDR (Incidente de Resolução de Demandas Repetitivas), por sua vez, também funciona como ferramenta de racionalização dos litígios de massa.

Não obstante os elementos de atratividade que a caverna algorítmica propaga (frequentemente interpretados como vantagens operacionais), subsistem nela riscos relevantes que demandam cautela. Primeiramente, a IA, por mais sofisticada que seja, apenas simula essas vivências, não sendo capaz de experimentá-las. Isso estabelece um limite ontológico intransponível entre máquinas e seres humanos, especialmente quando se trata de interpretar e aplicar o direito. A partir desse ponto, as decisões judiciais exigem mais do que a extração de padrões, pois demandam sensibilidade e compreensão do contexto.

Acresce a isso a opacidade intrínseca a determinados sistemas de inteligência artificial, como florestas aleatórias, máquinas de vetores de suporte (SVMs) e perceptrons multicamadas, cujas arquiteturas complexas dificultam a transparência dos critérios decisórios automatizados. Os vieses discriminatórios incorporados a esses agentes podem originar falhas tanto na programação do código quanto na alimentação dos sistemas de aprendizado automático, sobretudo quando submetidos a conjuntos de dados incompletos ou manipulados. Essa obscuridade fragiliza garantias constitucionais como o contraditório, a motivação e a ampla defesa, corroendo a confiança no juízo humano e comprometendo a legitimidade do próprio exercício jurisdicional.

No âmbito linguístico, a polissemia de determinadas palavras representa um desafio suplementar, visto que termos ambíguos demandam uma interpretação que atenda às especificidades do contexto. Ademais, a alienação cognitiva, decorrente da delegação do poder decisório à inteligência artificial, pode levar o operador do direito a perder o ímpeto de refletir, interpretar e problematizar, substituindo o exercício hermenêutico por respostas céleres e padronizadas. Além das implicações éticas e de responsabilidade de uma decisão influenciada por sistemas automatizados.

Com vistas a harmonizar inovação e prudência, o Conselho Nacional de Justiça instituiu, por meio da Resolução nº 332, de 21 de agosto de 2020, as primeiras diretrizes relativas à ética, transparência e

governança na criação e aplicação da inteligência artificial no âmbito do Poder Judiciário, além de outras disposições correlatas. Posteriormente, a Resolução nº 615, de 11 de março de 2025, aprimorou e ampliou tais diretrizes, regulamentando de forma mais detalhada o desenvolvimento e a governança das ferramentas de inteligência artificial. De modo particular, o Capítulo VI disciplinou o uso e a contratação de Modelos de Linguagem de Larga Escala (LLMs) e demais Sistemas de IA Generativa (IAGen) que criam conteúdos inéditos, diferente da IA explicativa que necessita de uma base de dados existente para suas funções.

A inteligência artificial e o processo civil, enquanto produtos do engenho humano, carregam consigo as marcas de uma racionalidade técnica. Assim, diante da relação entre processo judicial, decisão e tecnologia, quais os limites e possibilidades para se equilibrar os riscos e benefícios da inteligência artificial no campo do direito? A explicabilidade *post hoc* surge, assim, como ponte entre o funcionamento interno da máquina e a inteligibilidade para o operador humano, especialmente no âmbito da jurisdição. Através dela existem procedimentos visuais, textuais, por exemplo, por simplificação ou por verificação de relevância com o objetivo de tornar os processos decisórios das máquinas mais transparentes.

Entretanto, explicabilidade não é sinônimo de legitimidade, pois técnicas que aparentam clareza podem gerar “transparência simulada”, ocultando as verdadeiras estruturas de poder e os vieses embutidos nos algoritmos. A metáfora da *Matrix* é resgatada como símbolo de uma justiça automatizada, cuja lógica interna precisa ser decodificada para não aprisionar o direito em uma ilusão de eficiência. Então, como o “talvez” pode ser considerado uma resposta: talvez a explicabilidade *post-hoc* seja uma forma de ajudar a desvendar alguns mistérios de uma caixa de pandora, mas os verdadeiros limites e possibilidades repousam no próprio ser humano.

Enquanto a técnica for absolutizada e a consciência esquecida (juntamente com a filosofia e o próprio Direito) nenhuma regulamentação, tampouco mecanismos complementares, serão capazes de curar almas adoecidas. Há um fechamento do horizonte do pensar, um obscurecimento do ser. Trata-se a existência como um dado, ou seja, algo substantivo/objetificado que se converte em conceitos fixos, encobrindo o seu sentido originário. Ser é existência, visto que o ser humano se constitui como ser-no-mundo.

É entre os excessos da técnica (que impõem limites) e o silêncio da consciência (que guarda possibilidades) que se delineia a crise do nosso tempo. Ainda que múltiplas inteligências artificiais estejam presentes no Judiciário, a demanda processual permanece crescente, revelando que o cerne da problemática não reside na carência de ferramentas, mas na crise do próprio Estado e na ausência de sentido. Nesse desabrigo, a humanidade caminha em direção ao apagamento do ser que nela habita, visto que está enfeitiçada por algo que não possui a mesma forma de experienciar a vida, mas sim de simular e calcular (a IA não habita o mundo real).

REFERÊNCIAS

- ARISTÓTELES. *Ética a Nicômaco*. 4. ed. São Paulo: Nova Cultural, 1991.
- ARRIETA, Alejandro et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. **Information Fusion**, [s.l.], v. 58, 2019. Disponível em: <https://arxiv.org/abs/1910.10045>. Acesso em: 29 abril. 2025.
- ASHBY, William Ross. **An Introduction to Cybernetics**. Cybernetics, Chapman & Hall, London, 1956.
- BAPTISTA DA SILVA, Ovídio Araújo. **Jurisdição e execução na tradição romano-canônica**. 2. ed. rev. São Paulo: Revista dos Tribunais, 1997.
- BAPTISTA DA SILVA, Ovídio Araújo. **Processo e ideologia: o paradigma racionalista**. Rio de Janeiro: Forense, 2004.
- BATISTELLA, Paulo. IA Desafia Fronteiras do Direito e Exige Legislação “Viva”, avaliam especialistas. **Revista Conjur**. Disponível em: <https://www.conjur.com.br/2024-set-22/ia-desafia-fronteiras-do-direito-e-exige-legislacao-viva-avaliam-especialistas/>. Acesso em: 20 jun. 2024.
- BARROS, Vinicius Ribeiro. ODRs nos Conflitos Consumeristas e sua Obrigatoriedade. In: **Revista Esmat**. Palmas, v. 14, n. 24, p. 259 - 276, jul. à dez. 2022. DOI: 10.29327/270098.14.24.
- BENNETT, Max. **Uma Breve História da Inteligência : evolução, IA e os cinco avanços que nosso cérebro fez**. 1. ed. Rio de Janeiro: Alta Books, 2024.
- BÍBLIA SAGRADA. **Nova Tradução na Linguagem de Hoje**. Barueri (SP): Sociedade Bíblica do Brasil, 2009. 1280 p.
- BIONI, Bruno R. **Proteção de Dados Pessoais - A Função e os Limites do Consentimento**. 3. ed. Rio de Janeiro: Forense, 2021.
- BRASIL. Câmara dos Deputados. **Projeto de Lei da Câmara dos Deputados nº 21/2020**. Estabelece fundamentos, princípios e diretrizes para o desenvolvimento e a aplicação da inteligência artificial no Brasil; e dá outras providências. Disponível em: <https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2236340>. Acesso em:

20 jun. 2024.

BRASIL, Senado. **Projeto de Lei nº 2338/2023**. Dispõe sobre o uso da Inteligência Artificial. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/157233>. Acesso em 17 de julho de 2024.

BRASIL. **Lei n. 9.800**, de 26 de maio de 1999. Permite às partes a utilização de sistema de transmissão de dados para a prática de atos processuais. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/19800.htm. Acesso em: 22 de março. de 2022.

BRASIL. **Resolução nº 354**, de 02 de dezembro de 2020. Regulamenta a criação de soluções tecnológicas para a resolução de conflitos pelo Poder Judiciário por meio da conciliação e mediação. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3604>. Acesso em 03 de março de 2025.

BRASIL. **Senado Federal**. Senado aprova regulamentação da inteligência artificial; texto vai à Câmara. Brasília, DF: Agência Senado. Disponível em: <https://www12.senado.leg.br/noticias/materias/2024/12/10/senado-aprova-regulamentacao-da-inteligencia-artificial-texto-vai-a-camara>. Acesso em: 05 de jan. 2025.

BOEING, Daniel Henrique Arruda; ROSA, Alexandre Moraes da. **Ensinando um robô a julgar**: pragmática, discricionariedade, heurísticas e vieses no uso de aprendizado de máquina no judiciário. Florianópolis: Emais, 2020.

BOLESINA, Iuri.; GERVASONI, Tamiris A. Gamificação no Ensino Jurídico: uma experiência aplicada às gerações contemporâneas. In: Anais do IV Congresso Internacional uma Nova Pedagogia para a Sociedade Futura. | ISBN 978-85-68901-29-8.

BUTMAN, Judith.; ALLEGRI, Ricardo F. A Cognição Social e o Córtex Cerebral. In: **Psicologia: Reflexão e Crítica**. 2001, vol. 14, no. 2, pp. 275-279.

CAMARGO, Beatriz Meneghel Chagas. **A separação dos poderes e os freios e contrapesos na Constituição de 1988**: a atuação do poder judiciário. Ponta Grossa - PR: Atena, 2021.

CAMERON, James. **O Exterminador do Futuro**. Produção: Gale Anne Hurd. Estados Unidos: Orion Pictures, 1984. 1 DVD (107 min).

CAPPELLETTI, Mauro; GARTH, Bryant; TROCKER, Nicolò. **Rabels Zeitschrift für ausländisches und internationales Privatrecht/ The Rabel Journal of Comparative and International**

Private Law 40. Jahrg., H. 3/4, Der Schutz des Schwächeren im Recht, 1976, p. 669-717. Disponível em <http://www.jstor.org/579946967&uid=70&uid=4579946977&uid=2134>. Acesso em 04 de abril de 2024.

CAPPELLETTI, Mauro; GARTH, Bryant. **Acesso à Justiça**. Porto Alegre: Sergio Antonio Fabris Editor, 2002.

CASTELLS, Manuel. **A Sociedade em Rede**. Tradução Roneide Venâncio Majer. A Era da Informação: economia, sociedade e cultura. v.1. São Paulo: Paz e Terra, 1999.

CASTELLS, Manuel.; CARDOSO, Gustavo. **A Sociedade em Rede do Conhecimento à Acção Política**. Imprensa Nacional - Casa da Moeda: Lisboa, 2005.

CARBONERA, Joel; SOUZA, Clarisse de; GONÇALVES, Bernardo. O problema da explicação em Inteligência Artificial: considerações a partir da semiótica. Teccogs: **Revista Digital de Tecnologias Cognitivas**, TIDD | PUC-SP, São Paulo, n. 17, p. 59-75, jan-jun. 2018.

CARROLL, Lewis. **Alice's Adventures in Wonderland**. London: Macmillan, 1865.

CHEIN, Flávia. **Introdução aos Modelos de Regressão Linear**: um passo inicial para compreensão da econometria como uma ferramenta de avaliação de políticas públicas. Brasília: Enap, 2019.

COMMERFORD, Tim; ROCHA, Zack de la; MORELLO, Tom; WILK, Brad. Wake Up. In: **Rage Against the Machine**. Los Angeles: Epic Records, 1992

CONSELHO NACIONAL DE JUSTIÇA. **Justiça em Números 2024**: Barroso destaca aumento de 9,5% em novos processos. Agência CNJ de Notícias, Brasília, 28 de maio de 2024. Disponível Em: <https://www.cnj.jus.br/justica-em-numeros-2024-barroso-destaca-aumento-de-95-em-novos-processos/#:~:text=Quase%2084%20milh%C3%B5es%20de%20processos,servidores%20brasileiros%20para%20serem%20solucionados..> Acesso em 28 de fevereiro de 2025.

CONSELHO NACIONAL DE JUSTIÇA. **Pesquisa Uso de Inteligência Artificial IA no Poder Judiciário 2023**. Conselho Nacional de Justiça; Programa das Nações Unidas para o Desenvolvimento. Brasília: CNJ, 2024.

CONSELHO NACIONAL DE JUSTIÇA. **O Uso da Inteligência**

Artificial Generativa no Poder Judiciário Brasileiro: relatório de pesquisa. Brasília: CNJ, 2024.

CONSELHO NACIONAL DE JUSTIÇA. **Inscrições abertas para o XV Prêmio Conciliar é Legal.** Brasília: CNJ, 2024.

DAUGHERTY, Paulo; WILSON, James H. **Humano + Máquina:** reinventando o trabalho na era da IA. Tradução Wendy Campos. Rio de Janeiro, 2019.

DESCARTES, René. **Discurso do Método.** Tradução de Paulo Neves. Porto Alegre: L&PM, 2007.

DINIZ, Davi Monteiro. **Regulação da Inteligência Artificial:** iniciativas estaduais nos EUA. **Revista Conjur.** Disponível em: <https://www.conjur.com.br/2024-dez-17/regulacao-da-inteligencia-artificial-iniciativas-estaduais-nos-eua/>. Acesso em 20 de dezembro de 2024.

DONEDA, Danilo. Prefácio. In: BIONI, Bruno R. **Proteção de Dados Pessoais - A Função e os Limites do Consentimento.** 3. ed. Rio de Janeiro: Forense, 2021.

DWORKIN, Ronald. **Uma Questão de Princípio.** São Paulo: Martins Fontes, 2001.

ECHAI, Danielle Sales Espinoza. **Entre substancialismo e procedimentalismo:** elementos para uma teoria constitucional brasileira adequada à luz do paradigma neoconstitucionalista. Dissertação (Mestrado em Direito Público) - Programa de Pós-graduação em Direito, Universidade Federal de Alagoas (UFAL), Maceió, 2008.

ECO, Umberto. **Apocalípticos e Integrados.** São Paulo: Perspectiva, 1993, 5ª ed.

ESPÍNDOLA, Ângela Araújo da Silveira. Acesso à justiça reloaded: decodificando a matrix para refundar a jurisdição. In: SALDANHA, Alexandre Henrique Tavares; SALDANHA, Paloma Mendes; PIMENTEL, Alexandre Freire. **Estudo em Direito Digital.** 1.ed. Recife: FASA, 2022.

ESPÍNDOLA, Ângela Araújo da Silveira; SILVA, Everton Luís da. Ativismo Judicial e a “Não-Resposta” como Possibilidade: o papel do “talvez” na refundação da jurisdição judicial. **Revista Jurídica (FURB).** v. 22, no. 49, set./dez. 2018.

ESPÍNDOLA, Ângela Araújo da Silveira; FANTONELLI, Miliane dos Santos. Decisão judicial, vozes citadas e vozes sobrepostas: é audível a

polifonia processual? **Revista do Instituto de Hermenêutica Jurídica**. Belo Horizonte, v.17, n.26, jul./dez. 2019.

FARIA, Rodrigo Martins. **Plataformas de Resolução de Disputas On-Line Orientadas Por Inteligência Artificial**. s.d. Disponível em: <https://www.tjmg.jus.br/lumis/portal/file/fileDownload.jsp?fileId=8ACC82D28C3C784A018C4437BAB568E3>. Acesso em 20 de março de 2025.

FARIA, Rodrigo Martins. Os Núcleos de Justiça 4.0 como Instrumentos de Cooperação Judiciária para a Gestão de Demandas Repetitivas. **Revista Eletrônica dos Grupos de Estudos da EJEJF**. 2022. ISSN: 2764-6742.

FILHO, Eduardo Isaia; LINDEMANN, Vanessa. **Algoritmos e Programação I**. Canoas: Ed. ULBRA, 2013. 144p.

FILHO, Elmano Rodrigues Pinheiro. Regulação da inteligência artificial: reflexões sobre a definição de uma autoridade competente no Brasil. In: **Inteligência artificial e Sociedade Conectada**. 1. ed. Rio de Janeiro: ITS - Instituto de Tecnologia e Sociedade, 2025.

FUNDAÇÃO GETULIO VARGAS. **Online Dispute Resolution**: um estudo à luz da resolução no 358 do CNJ. Rio de Janeiro: FGV, 2023.

GAARDER, Jostein. **O Mundo de Sofia**. Tradução: Catarina Belo. Baixo: Editorial Presença, 1995. 461 p.

GADAMER, Hans-Georg. **Verdade e Método**. Tradução de Flávio Paulo Meurer. 3. ed. Petrópolis: Vozes, 1999.

GALANTER, Marc. **Por que “quem tem” sai na frente**: especulações sobre os limites da transformação no direito. São Paulo : FGV Direito SP, 2018.150 p.

HEIDEGGER, Martin. **Ser e Tempo**. 15. ed. Tradução de Márcia de Sá Cavalcante. Petrópolis: Vozes, 2005a.

HOFFMAM, Fernando; MORAIS, Jose Luis Bolzan de. Fundamentos Hermenêuticos para um Processo Civil Constitucionalmente Adequado. **Revista da Faculdade de Direito-RFD-UERJ** - Rio de Janeiro, n. 30 , p. 266- 296, dez. 2016.

INSTITUTO DE TECNOLOGIA E SOCIEDADE DO RIO. **Relatório Matriz comparada de obrigações**: PL 2.338/2023 vs. EU AI Act. Rio de Janeiro, 2023. Disponível em: <https://itsrio.org/pt/publicacoes/relatorio-matriz-comparada-de-obrigacoes-pl-2338-2023-vs->

eu-ai-act-2/. Acesso em: 14 fev. 2025.

INWOOD, Michael. **Dicionário de Heidegger**. Tradução de Luísa Buarque de Holanda. Rio de Janeiro: Jorge Zahar, 2002.

ISAIA, Cristiano Becker. **Processo civil e hermenêutica: a crise do procedimento ordinário e o redesenhar da jurisdição processual civil pela sentença (democrática) liminar de mérito**. Curitiba: Juruá, 2012.

ISAIA, Cristiano Becker. Os Desafios da Jurisdição Processual Civil no Século 21. In: TYBUSCH, Jerônimo Siqueira [et al.]. **Direitos Emergentes na Sociedade Global: anuário do programa de pós-graduação em direito da UFSM**. Unijuí, 2013.

ISAIA, Cristiano Becker. **A Necessidade de Uma Compreensão Hermenêutica e Democrática do Direito Processual Civil e o Desvelar do Caso Concreto com a Sentença Liminar de Mérito: a jurisdição-processual herdada e a jurisdição-processual a ser desenhada por uma filosofia no processo**. Tese de Doutorado Disponível em: <http://repositorio.jesuita.org.br/handle/UNISINOS/3996>. Acesso em: 01 de abril de 2024.

JACOBSEN, Gabriel. Produção de decisões e análises de processos: como a inteligência artificial é utilizada na Justiça do RS. **GZH**, Porto Alegre, 15 jul. 2025. Disponível em: <https://gauchazh.clicrbs.com.br/geral/noticia/2025/07/producao-de-decisoes-e-analises-de-processos-como-a-inteligencia-artificial-e-utilizada-na-justica-do-rs-cmd3lidjt003e014m9yg13arg.html>. Acesso em: 15 jul. 2025.

JIAWEI, Gu. et al. A Survey on LLM-as-a-Judge. **arXiv**, v. 1, n. 1. Cornell University, 2025. Disponível em: <https://doi.org/10.48550/arXiv.2411.15594>; Acesso em 14 de março de 2025.

KAHLMAYER-MERTENS, Roberto S. **10 Lições sobre Gadamer**. Petrópolis, RJ: Vozes, 2017.

KAHNEMAN, Daniel. **Rápido e Devagar: duas formas de pensar**. Rio de Janeiro: Objetiva, 2012.

KARDEC, Allan. **O Livro dos Espíritos**. Tradução de Salvador Gentile. 183. ed. São Paulo: Araras, 2021.

KISSINGER, Henry A.; SCHMIDT, Eric; HUTTENLOCHER, Daniel. **A Era da IA**. Rio de Janeiro: Alta Books, 2023.

KRANTZ, Tom; JONKER, Alexandra. **What is the AI Bill of Rights?**. IBM, 27 de setembro de 2024. Pensar. Disponível em: <https://www.ibm.com>.

com/think/topics/ai-bill-of-rights. Acesso em: 07 de janeiro de 2025.

LAGE, Fernanda de Carvalho. **Inteligência Artificial na Repercussão Geral**: análise e proposições da vanguarda de inovação tecnológica no Poder Judiciário brasileiro. Tese de Doutorado. Programa de Pós-graduação em Direito da Universidade de Brasília, Brasília, 2020.

LEE, Kai-Fu. **Inteligência Artificial**: como os robôs estão mudando o mundo, a forma como amamos, nos relacionamos, trabalhamos e vivemos. Tradução de Marcelo Barbão. 1. ed. - Rio de Janeiro: Globo Livros, 2019.

LEONE, Priscilla Novaes. Admirável Chip Novo. In: **Admirável Chip Novo**. São Paulo: Deckdisc, 2003.

LIMA, Daniel Henrique Sprotte. **Da Cultura do Litígio à do Consenso**: o uso de online dispute resolution na Comarca De Araquari (SC). Dissertação de Mestrado. Programa de Pós-Graduação em Direito da Universidade Federal de Santa Catarina Florianópolis, 2019.

LUCHIARI Valeria Ferioli Lagrasta. **Mediação Judicial**: Análise da Realidade Brasileira: Origem e Evolução até a Resolução n. 125 do Conselho Nacional de Justiça Coordenadores: Ada Pellegrini Grinover, Kazuo Watanabe. Rio de Janeiro: Forense, 2012.

LUGER, George F. **Inteligência Artificial**. Tradução Daniel Vieira; Revisão técnica Andréa Iabrudi Tavares. 6. ed. São Paulo: Pearson Education do Brasil, 2013.

MARQUES, Júlio V. M. et al. **Explorando a explicabilidade da inteligência artificial**: técnicas para compreender e interpretar modelos de aprendizado de máquina. Disponível: <https://books-sol.sbc.org.br/index.php/sbc/catalog/download/132/574/886?inline=1>. Acesso em 05 de abril de 2025.

MAVRAKIS, Nicolás. **Heidegger e o “esquecimento do Ser”**: um vigoroso pensamento contra a civilização tecnificada. Instituto Humanitas Unisinos, 2021. Disponível em: <https://ihu.unisinos.br/categorias/609611-heidegger-e-o-esquecimento-do-ser-um-vigoroso-pensamento-contra-a-civilizacao-tecnificada>. Acesso em 11 de abril de 2025.

MCCARTHY, John. From here to human-level AI. In: **Artificial Intelligence**. v. 171. 18. ed. Elsevier BV, 2007. ISSN: 0004-3702. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0004370207001476>. Acesso em: 15 de setembro de 2024.

MCCULLOCH, Warren Sturgis; PITTS, Walter. **A Logical Calculus of the Ideas Immanent in Nervous Activity**. *Bulletin of Mathematical Biology*, v. 52, n. 1/2, p. 99-115, 1990.

MELO, Jairo. **Inteligência Artificial no Poder Judiciário Brasileiro**. Tribunal de Justiça do Distrito Federal e dos Territórios – TJDF, 2023.

MENDES, Aluisio Gonçalves de Castro. **Incidente de resolução de demandas repetitivas**: sistematização, análise e interpretação do novo instituto processual. Rio de Janeiro: Forense, 2017.

MENDES, Aluisio Gonçalves de Castro; SILVA, Larissa Clare Pochmann da. Acesso à justiça: uma releitura da obra de Mauro Capelletti e Bryant Garth a partir do Brasil 40 anos depois. **Revista Quaestio Iuris**, [S. l.], v. 8, n. 03, p. 1827–1858, 2015. DOI: 10.12957/rqi.2015.19385. Disponível em: <https://www.e-publicacoes.uerj.br/quaestioiuris/article/view/19385>. Acesso em: 04 de abril de 2024.

MEINKE, Alexander et al. **Frontier Models are Capable of In-context Scheming**. Apollo Research, 2025. Disponível: <https://arxiv.org/pdf/2412.04984>. Acesso em 04 de maio de 2025.

MOMOLLI, Andréia. **Hermenêutica Jurídica e Inteligência Artificial no Processo Jurisdicional da Sociedade em Rede**. Dissertação (Mestrado em Direito) - Programa de Pós-graduação em Direito, Universidade Federal de Santa Maria (UFSM), Santa Maria, 2020.

MONTEIRO, António Pinto. “Qui facit per alium, facit per se” — será ainda assim na era da robótica? In: _____. (org.). **Direito e Robótica**. Faculdade de Direito da Universidade de Coimbra, 2020.

MINSKY, Marvin. **Steps Toward Artificial Intelligence**. Proceedings of the IRE, v. 49, n. 1, p. 8-30, 1961.

MINSKY, Marvin. **The Society of Mind**. Nova Iorque: Simon e Schuster, 1986.

MUNDIM, Luís Gustavo Reis. Inteligência Artificial e Precedentes: a governamentalidade algorítmica e o neoliberalismo processual como entraves ao devido processo. **Rev. Novas Conject.** Fac. Direito UEMS, Paranaíba. v. 1, n. 1. jan./dez. 2025.

NEWELL, Allen. **Intellectual Issues in the History of Artificial Intelligence**. In: MACHLUP, Fritz; MANSFIELD, Una (Eds.). *The Study of Information: Interdisciplinary Messages*. New York: John Wiley & Sons, 1983.

NORTHFLEET, Ellen Gracie. **A utilização do fax no poder judiciário.** Revista dos Tribunais. São Paulo, n.728, p.122-127, jun.1996.

NUNES, Dierle José Coelho. **Comparticipação e Policentrismo:** horizontes para democratização processual civil. 2008. Tese (Doutorado em Direito Processual) - Programa de Pós-graduação em Direito, Pontifícia Universidade Católica de Minas Gerais (PUC Minas), Belo Horizonte, 2008.

NUNES, Dierle José Coelho. Processo Civil, Vieses Cognitivos e Tecnologia: alguns desafios. In LUCON, Paulo Henrique dos (Org.); OLIVEIRA, Pedro Miranda de (Org.). **Panorama atual do novo CPC.** São Paulo: Tirant lo Blanch, 2019.

NUNES, Dierle José Coelho. Virada Tecnológica no direito processual: fusão de Justiça Digital e Inteligência Artificial: conhecimentos para geração de uma nova justiça centrada no ser humano. In: **Revista de Processo**, vol. 344, p. 403-429, Out.2023a.

NUNES, Dierle José Coelho. Uma Introdução sobre o Uso dos Modelos de Inteligência Artificial Analíticas e Generativas no Direito Processual. In: VADELL, Lorenzo M. Bujosa. **O Sistema Processual do Século XXI:** novos desafios. Londrina: Thoth, 2023b.

NUNES, Dierle José Coelho; ANDRADE, Otávio Morato de. O Uso da Inteligência Artificial Explicável enquanto Ferramenta para Compreender Decisões Automatizadas: possível caminho para aumentar a legitimidade e confiabilidade dos modelos algorítmicos?. **Revista Eletrônica do Curso de Direito da UFSM**, [S. l.], v. 18, n. 1, p. e69329, 2023c. DOI: 10.5902/1981369469329.

NUNES, Dierle; LUD, Natanael; PEDRON, Flávio Quinaud. **Desconfiando da Imparcialidade dos sujeitos processuais:** um estudo sobre os vieses cognitivos, a mitigação de seus efeitos e o debiasing, 2a. ed. Salvador: JusPodivm, 2020.

NUNES, Dierle; MORATO, Otávio. A Explicabilidade da Inteligência Artificial e o Devido Processo Tecnológico. **Revista Conjur.** Disponível em: <https://www.conjur.com.br/2021-jul-07/opinioao-explicabilidade-ia-devido-processo-tecnologico/>. Acesso em 10 de abril de 2025.

NUNES, Dierle; PAOLINELI, Camila. Acesso à Justiça e virada tecnológica no sistema de Justiça brasileiro: gestão tecnológica de disputas e o alinhamento de expectativas para uma transformação com foco no cidadão – novos designs, arquitetura de escolhas e tratamento adequado de disputas. In: NUNES, Dierle [et all] **Direito Processual e**

Tecnologia: os impactos da virada tecnológica no âmbito mundial.

NUNES, Dierle; RODRIGUES, Larissa H. A. O contraditório e sua implementação pelo design: design thinking, legal design e visual law como abordagens de implementação efetiva da influência. In: NUNES, Dierle; LUCON, Paulo Henrique dos Santos; WOLKART, Erik Navarro (Coord.). **Inteligência artificial e direito processual: os impactos da virada tecnológica no direito processual**. Salvador: JusPodivm, 2020.

OLIVEIRA, Natália. Gibberlink Mode: A Nova Linguagem da IA Que Levanta Questionamentos Sobre Transparência. **Itshow**, 07 de março de 2025. Disponível em: <https://itshow.com.br/gibberlink-mode-comunicacao-entre-ia/>. Acesso em 09 de abril de 2025.

OLIVEIRA, Luiz Fernando de. Transmissão Sináptica. **Revista Brasileira de Anestesiologia**. Vol. 44 : Nº 1, Janeiro - Fevereiro, 1994.

O'NEIL, Cathy. **Algoritmos de Destruição em Massa:** como o *big data* aumenta a desigualdade e ameaça a democracia. Santo André: Rua do Sabão, 2020.

PASQUALE, Frank. Normative Dimensions of Consensual Application of Black Box Artificial. In: FURBINO, Meire.(Org.); BOCCHINO, Lavínia Assis.(Org.); LIMA, Maria Jocélia Nogueira (Org.) **A Inteligência Artificial:** a (des)serviço do estado de direito. Belo Horizonte: CAPES: Programa de Pós-graduação em Direito-PUC Minas: RTM, 2023.

PEDRON, Flávio; MENEZES, Rafael; TORRES, Tiago Henrique. Um “novo normal” para o novo judiciário? Intersecções entre Robinson Crusoe e os diferentes litigantes no acesso à justiça durante (e pós) a Pandemia. **Revista Eletrônica de Direito Processual**, v. 23, n. 2, 2022.

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

PEREIRA, Roberto Horácio. **Qualia**. Compêndio em Linha de Problemas de Filosofia Analítica. Campo Grande, 2013. ISBN: 978-989-8553-22-5

PINHEIRO, Patricia Peck. **Direito Digital**. 7. ed. São Paulo: Saraiva Educação, 2021.

PINTO, Henrique Alves. A utilização da inteligência artificial no processo de tomada de decisões: por uma necessária accountability. **Revista de Informação Legislativa:** RIL, Brasília, DF, v. 57, n. 225, p.

43-60, jan./mar. 2020. Disponível em: http://www12.senado.leg.br/ril/edicoes/57/225/ril_v57_n225_p43

PREMUZIC, Tomas Chamorro. **Eu, Humano: IA, automação e a busca para recuperar o que nos torna únicos**. Traduzido por Eveline Machado. Rio de Janeiro: Alta Books, 2024.

RIBEIRO, Marcelo. **Processo Civil**. 3. ed. Rio de Janeiro: Método, 2023.

RUSSELL, Stuart J.; NORVIG, Peter. **Inteligência artificial**. tradução Regina Célia Simille. Rio de Janeiro: Elsevier, 2013.

SADIN, Éric. O tecnoliberalismo lança-se à Conquista Integral da Vida: entrevista com Éric Sadin. **Instituto Humanitas Unisinos – IHU**, São Leopoldo, 24 abr. 2017. Disponível em: <https://www.ihu.unisinos.br/categorias/186-noticias-2017/568991-o-tecnoliberalismo-lanca-se-a-conquista-integral-da-vida-entrevista-com-eric-sadin>. Acesso em: 9 maio 2025.

SALE, Ana Débora Rocha. **Inteligência Artificial (IA) à Luz da Teoria da Decisão: um estudo sobre a utilização da IA em decisões judiciais**. São Paulo: Editora Dialética, 2023.

SALOMÃO, Luis Felipe (coord.). **Marco legal da Inteligência Artificial: nota técnica sobre o Projeto de Lei 21/2020**. Rio de Janeiro: FGV conhecimento – Centro de inovação, administração e pesquisa do Judiciário, 2021. Disponível em: https://ciapj.fgv.br/sites/ciapj.fgv.br/files/ciapj_fgv_notatecnica_ia.pdf Acesso em 06 out. 2021.

SALZBERG, Steven L. Book Review: **C4.5: Programs for Machine Learning**, de J. Ross Quinlan. Machine Learning, v. 16, p. 235-240, 1994.

SCHMIDT, Eric; HUTTENLOCHER, Daniel; KISSINGER, Henry A. **A Era da IA: e nosso futuro como humanos**. Traduzido por Vanessa Schreiner. Rio de Janeiro: Alta Books, 2023.

SCOTT, Ridley. **Blade Runner**. Produção: Michael Deeley. Estados Unidos: Warner Bros., 1982. 1 DVD (117 min).

SILVEIRA, Ricardo Geraldo R. **Acesso à Justiça: o direito fundamental em um ambiente de recursos escassos**. São Paulo: Almedina, 2020.

SIMON, Herbert Alexander. **The Sciences of the Artificial**. 3. ed. Cambridge: MIT Press, 1996.

STRECK, Lenio Luiz. **Hermenêutica jurídica e(m) crise:** uma exploração hermenêutica da construção do Direito. 11. ed. rev., atual. e ampl. – Porto Alegre: Livraria do Advogado Editora, 2014.

STRECK, Lenio Luiz. Apresentação. In: **O Impacto das Novas Tecnologias nos Direitos Fundamentais**. Joaçaba: Unoesc, 2015.

STRECK, Lenio Luiz. O(a) Robô Maria, Precedentes e o Direito a um Julgamento Humano. **Revista Conjur**. Disponível: <https://www.conjur.com.br/2025-fev-13/oa-robo-maria-precedentes-e-o-direito-a-um-julgamento-humano/>. Acesso em 11 de abril, de 2025.

STRECK, Lenio Luiz.; MORAIS, Jose Luis Bolzan de. **Ciência política e teoria do estado**. 8. ed. rev. e atual. – Porto Alegre: Livraria do Advogado Editora, 2014.

STRECK, Lênio Luiz; BERNSTES, Luísa Giuliani; GOMES, Jefferson de Carvalho. Inteligência Artificial: mesmos problemas, mas na versão hi-tech. Constituição, Economia e Desenvolvimento: **Revista Eletrônica da Academia Brasileira de Direito Constitucional**, [S. l.], v. 13, n. 25, p. 333–342, 2021.

STRECK, L. L.; JUNG, L. Hermenêutica e Inteligência Artificial: Por uma Alternativa Paradigmática ao Imaginário Técnico-Jurídico. **Direito Público**, [S. l.], v. 21, n. 110, 2024. DOI: 10.11117/rdp.v21i110.7689. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/7689>. Acesso em: 5 ago. 2024.

TRIBUNAL REGIONAL FEDERAL DA 5ª REGIÃO – TRF5. **Levantamento quantitativo de IRDRs Processados e Julgados entre 2016 e 1º de outubro de 2024**. 2024. Disponível em: https://arquivos.trf5.jus.br/TRF5/NUGEPNAC/2024/10/01/20241001_BA040A_Levantamento_do_quantitativo_de_IRDRs_de_2016_a_2024_atualizado_ate_01_de_outubro_de_2024.PDF. Acesso em: 18 de mar. de 2025.

TEIXEIRA, Gujo. [Foto]. Instagram, 03 de abril de 2025. Disponível: <https://www.instagram.com/p/DH-1X6ux1xX/?igsh=MXhjbTBldmdqbnNpag==>. Acesso em: 03 abr. 2025.

TRIBUNAL DE JUSTIÇA DO ESTADO DO PARÁ – TJPA. **IRDR e Incidente de Assunção de Competência**. 2021. Disponível em: <https://www.tjpa.jus.br/CMSPortal/VisualizarArquivo?idArquivo=982113>. Acesso em: 18 de mar. de 2025.

TURING, Alan Mathison. **On Computable Numbers, with an**

Application to the Entscheidungsproblem. Proceedings of the London Mathematical Society, v. 2, n. 42, p. 230-265, 1936.

TYBUSCH, Jerônimo. Tempo e Possibilidades. **Diário de Santa Maria**, Santa Maria, 31 mar. 2025.

TYBUSCH, Jerônimo. Técnica e Humanidade. **Diário de Santa Maria**, Santa Maria, 2025.

UNESCO. **Diretrizes da UNESCO para o Uso de Sistemas de IA em Tribunais.** Disponível em: https://unesdoc.unesco.org/ark:/48223/pf0000390781_por. Acesso em 07 de abr. de 2025.

UNIÃO EUROPEIA. **Regulamento da Inteligência Artificial.** Disponível em: <https://www.consilium.europa.eu/pt/policies/artificial-intelligence/#0>. Acesso em 05 de janeiro de 2025.

UNIÃO EUROPEIA. **Regulamento do Parlamento Europeu e do Conselho.** Disponível em: <https://eur-lex.europa.eu/legal-content/PT/ALL/?uri=CELEX:32024R1689>. Acesso em 04 de janeiro de 2025.

UNIVERSIDADE FEDERAL DE SANTA MARIA. **FORTRAN:** Um dinossauro das linguagens de programação volta depois de 10 anos. Disponível em: <https://www.ufsm.br/pet/sistemas-de-informacao/2021/05/04/fortran-um-dinossauro-das-linguagens-de-programacao-volta-depois-de-10-anos>. Acesso em 20 de fevereiro de 2025.

VIDAL, Ana Carolina; SOUSA, Marcelo Sequeira de. **UE é Pioneira na Regulamentação de Sistemas de Inteligência Artificial.** Jota, UE, 04 de novembro de 2024. RegulaEuro. Disponível em: <https://www.jota.info/opiniao-e-analise/colunas/regula-euro/ue-e-pioneira-na-regulamentacao-de-sistemas-de-inteligencia-artificial>. Acesso em: 06 de janeiro de 2024.

VILAÇA, Gracila. **Matrix e o projeto neocolonial do Vale do Silício.** Nexo Ensaio, Brasil, 21 de janeiro de 2025. Disponível em: <https://www.nexojornal.com.br/matrix-mark-zuckerberg-neocolonialismo>. Acesso em 23 de janeiro de 2024.

VÁZQUEZ, Favio. **Deep learning made easy with Deep Cognition.** 2017. Disponível em: <https://becominghuman.ai/deep-learning-made-easy-with-deep-cognition-403fbe445351>. Acesso em: 29 de set. de 2024.

VIRILIO, Paul. **O Espaço Crítico.** Tradução Paulo Roberto Pires. Rio de Janeiro: Ed.34, 1993.

WACHOWSKI, Lana; WACHOWSKI, Lilly (Direção e Roteiro).

Matrix. Produção: Joel Silver. Estados Unidos: Warner Bros. Pictures, 1999. 1 DVD (136 min).

WARAT, Luís Alberto. **O Direito e sua Linguagem.** Porto Alegre, safe, 2000.

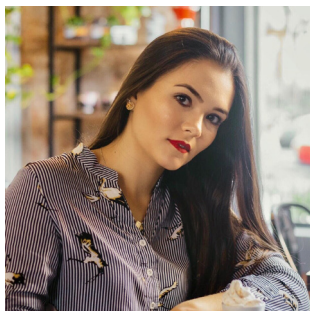
WEBB, Amy. **Os Nove Titãs da IA.** 1.ed. Rio de Janeiro: Alta Books, 2020.

WIENER, Norbert. **Cibernética e Sociedade:** o uso humano de seres humanos. Tradução de José Paulo Paes. 2. ed. São Paulo: Cultrix, 1954.

ZAMBERLAN, Alexandre. et al. Técnicas de Inteligência Artificial Aplicadas ao Direito: representação de conhecimento e raciocínio automatizado. In: **XXV Simpósio de Ensino, Pesquisa e Extensão.** ISSN: 2316-9745. Disponível em: <http://doi.org/10.48195/sepe2021-235>. Acesso em: 15 de julho de 2024.

ZENG, Yi; WANG, Ling. **Fei-Fei Li:** Artificial Intelligence is on its way to reshape the world, National Science Review, Volume 4, Issue 3, May 2017, Pages 490–492.

SOBRE A AUTORA



CAROLINE DA ROSA CAVALHEIRO

Mestra em Direito pelo Programa de Pós-Graduação em Direito da Universidade Federal de Santa Maria (PPGDUFMSM), na linha de pesquisa Direitos da Sociedade em Rede: atores, fatores e processo na mundialização. Pós-graduada no nível de especialização em Direito e Processo do Trabalho pela UniRitter. Graduada em Direito pela Universidade Luterana do Brasil - Campus Santa Maria/RS - (Bolsista PROUNI) (2022). Apresentou e publicou trabalhos em eventos nacionais e internacionais. Atualmente, desde 2020, atua como colaboradora da Ulbra/SM.

Resultado de uma pesquisa rigorosa e interdisciplinar, esta obra investiga criticamente as transformações do Direito Processual diante da incorporação crescente da inteligência artificial, dos sistemas de comportamento inteligente e da lógica tecnoneoliberal que atravessa as instituições judiciais contemporâneas. Fruto de Dissertação de Mestrado desenvolvida no Programa de Pós-Graduação em Direito da Universidade Federal de Santa Maria, o livro articula Direito, filosofia, ciência da computação, teoria do processo e estudos críticos da tecnologia para analisar como algoritmos, automação e plataformas digitais vêm reconfigurando a tomada de decisão judicial. Distanciando-se de abordagens tecnocráticas ou entusiastas, a autora enfrenta os limites, riscos e tensões inerentes à algoritmização da justiça, examinando o funcionamento técnico dos sistemas automatizados, seus vieses potenciais e suas implicações normativas, éticas e constitucionais. Com precisão conceitual e linguagem acessível, a obra traduz conceitos complexos do campo computacional sem abrir mão da densidade teórica. Ao problematizar o processualismo tecnoneoliberal e sua lógica gerencialista de padronização e aceleração decisória, o livro convida o leitor a refletir sobre os impactos estruturais da tecnologia no Direito e sobre a necessidade de preservar a racionalidade jurídica, a integridade interpretativa e os fundamentos democráticos do processo. Trata-se de uma leitura indispensável para quem busca compreender, com profundidade crítica, os desafios do Judiciário na era da inteligência artificial.

ISBN 978-656135214-7



EDITORA

ILUSTRAÇÃO