

UNIVERSIDADE FEDERAL DE OURO PRETO
Departamento de Direito

Gabriel Filipe Ferreira Gomes

**AUTOMATIZAÇÃO DE DECISÕES JUDICIAIS: Potencialidades e Limites Éticos da
Utilização de IAs Generativas no Processo Decisório**

Ouro Preto
2025

Gabriel Filipe Ferreira Gomes

**AUTOMATIZAÇÃO DE DECISÕES JUDICIAIS: Potencialidades e Limites Éticos da
Utilização de IAs Generativas no Processo Decisório**

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

Orientador: Prof.^a Dra. Juliana Evangelista de
Almeida

Ouro Preto

2025



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE OURO PRETO
REITORIA
ESCOLA DE DIREITO, TURISMO E MUSEOLOGIA
DEPARTAMENTO DE DIREITO



FOLHA DE APROVAÇÃO

Gabriel Filipe Ferreira Gomes

AUTOMATIZAÇÃO DE DECISÕES JUDICIAIS: Potencialidades e Limites Éticos da Utilização de IAs Generativas no Processo Decisório

Monografia apresentada ao Curso de Direito da Universidade Federal de Ouro Preto como requisito parcial para obtenção do título de Bacharel em Direito

Aprovada em 03 de Setembro de 2025

Membros da banca

Doutora - Juliana Evangelista de Almeida - Orientador(a) - Universidade Federal de Ouro Preto

Doutora - Beatriz Schettini - Universidade Federal de Ouro Preto

Mestrando - Vinícios Pereira Teixeira - Universidade Federal de Ouro Preto

Juliana Evangelista de Almeida, orientadora do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 03/09/2025



Documento assinado eletronicamente por **Juliana Evangelista de Almeida, PROFESSOR DE MAGISTERIO SUPERIOR**, em 04/09/2025, às 15:19, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0973395** e o código CRC **2AD853EB**.

Aos meus pais, que com coragem silenciosa e amor incansável me ensinaram a caminhar, e cujo suor derramado no chão hoje guia cada um dos meus passos.

AGRADECIMENTOS

Serei breve — não por economia de sentimentos, mas porque há gratidões que a palavra não alcança. O que aqui escrevo grava o papel e repousa no tempo, mas é na memória dos gestos, no brilho dos olhares e nos silêncios partilhados que a vida verdadeiramente agradece. O que marca o coração, quase nunca cabe inteiro na escrita.

Agradeço, em primeiro lugar, a Deus, pela força que guiou meus passos, mesmo nos momentos em que eu me sentia perdido.

Aos meus pais, Nilson e Sandra, que são terra firme e abrigo. Suas mãos me ensinaram o valor da luta, e seus olhos, o valor da ternura. Em cada sacrifício e em cada gesto de cuidado, vocês semearam em mim a coragem de seguir.

À minha namorada, Laís, agradeço não só pelo apoio constante, mas também pela luz que trouxe aos meus dias. A felicidade que você me dá não é só conforto — é centelha criadora e se fez também inspiração para este trabalho. Você foi poesia enquanto tudo era apenas esforço e cansaço.

Aos meus familiares, aos poucos amigos, e aos professores que passaram pelo meu caminho, agradeço por terem contribuído, cada um à sua maneira, para que este trabalho se tornasse possível.

“O homem está condenado a ser livre. Condenado porque não se criou a si próprio; e, no entanto, livre porque, uma vez lançado ao mundo, é responsável por tudo quanto fizer”.

(Jean-Paul Sartre)

RESUMO

O presente trabalho buscou investigar os limites ético-jurídicos da utilização de Inteligências Artificiais Generativas na redação de decisões judiciais, à luz das exigências constitucionais do Estado Democrático de Direito. A pesquisa parte do reconhecimento de que os modelos generativos de linguagem, apesar de sua fluência textual e aparente coerência argumentativa, operam segundo lógica estatística autorregressiva, sem genuína compreensão de significados, responsabilidade epistêmica ou abertura hermenêutica ao dissenso. Mediante análise crítica da arquitetura técnica dos *Large Language Models*, com destaque para a arquitetura *Transformer* e os mecanismos de treinamento supervisionado e retroalimentação por reforço, examina-se a desconexão entre a lógica de correlação algorítmica e o ideal hermenêutico-dialético da jurisdição. O estudo demonstra que a automatização da decisão judicial, ao suprimir o juízo hermenêutico e os elementos compreendidos na fundamentação constitucionalmente exigida, implica riscos à legitimidade democrática do processo, à inteligibilidade da decisão e à integridade do Direito. A metodologia adotada é qualitativa, de cunho teórico-bibliográfico, com abordagem jurídico-dogmática e base interdisciplinar. Conclui-se que, embora as IAGs possam assumir funções auxiliares na sistematização da linguagem jurídica, sua utilização no núcleo da atividade decisória revela-se incompatível com os pressupostos normativos e principiológicos do Estado Democrático de Direito.

Palavras-chave: Inteligência Artificial Generativa; Decisão Judicial; Fundamentação; Opacidade Algorítmica; Jurisdição; Large Language Models.

ABSTRACT

The present work sought to investigate the ethical-legal limits of the use of Generative Artificial Intelligences in the drafting of judicial decisions, in light of the constitutional requirements of the Democratic Rule of Law. The research starts from the recognition that generative language models, despite their textual fluency and apparent argumentative coherence, operate according to an autoregressive statistical logic, without genuine understanding of meanings, epistemic responsibility, or hermeneutic openness to dissent. Through critical analysis of the technical architecture of Large Language Models, with emphasis on the Transformer architecture and the mechanisms of supervised training and reinforcement learning, the disconnection between the logic of algorithmic correlation and the hermeneutic-dialectical ideal of jurisdiction is examined. The study demonstrates that the automation of judicial decision-making, by suppressing hermeneutic judgment and the elements comprised in the constitutionally required reasoning, implies risks to the democratic legitimacy of the process, the intelligibility of the decision, and the integrity of Law. The methodology adopted is qualitative, of a theoretical-bibliographic nature, with a legal-dogmatic approach and an interdisciplinary basis. It is concluded that, although GAI's may assume auxiliary functions in the systematization of legal language, their use in the core of the decision-making activity proves to be incompatible with the normative and principled assumptions of the Democratic Rule of Law.

Keywords: Generative Artificial Intelligence; Judicial Decision; Legal Reasoning; Algorithmic Opacity; Jurisdiction; Large Language Models.

LISTA DE ABREVIATURAS E SIGLAS

BERT	<i>Bidirectional Encoder Representations from Transformers</i>
CNJ	Conselho Nacional de Justiça
COMPAS	<i>Correctional Offender Management Profiling for Alternative Sanctions</i>
CPC	Código de Processo Civil
DNN	<i>Deep Neural Network</i>
FFN	<i>Feed Forward Network</i>
GAN	<i>Generative Artificial Intelligence</i>
GELU	<i>Gaussian Error Linear Unit</i>
GenAI	<i>Generative Artificial Intelligence</i>
GPT	<i>Generative Pre-trained Transformer</i>
IA	Inteligência Artificial
IAG	Inteligência Artificial Generativa
LIME	<i>Local Interpretable Model-Agnostic Explanations</i>
LLaMA	<i>Large Language Model Meta AI</i>
LLM	<i>Large Language Models</i>
LSTM	<i>Long Short-Term Memory</i>
NLP	<i>Natural Language Processing</i>
ReLU	<i>Rectified Linear Unit</i>
RLHF	<i>Reinforcement Learning from Human Feedback</i>
RNN	<i>Recurrent Neural Network</i>
SFT	<i>Supervised Fine-Tuning</i>
SHAP	<i>Shapley Additive Explanations</i>
STF	Supremo Tribunal Federal
XAI	<i>Explainable Artificial Intelligence</i>

LISTA DE ILUSTRAÇÕES

Figura 1 – Estrutura de camadas e operação da arquitetura <i>Transformer Decoder-Only</i>	13
--	----

SUMÁRIO

1. INTRODUÇÃO.....	9
2. A INTELIGÊNCIA ARTIFICIAL GENERATIVA.....	11
2.1. Fundamentos Conceituais e Paradigmas Tecnológicos das IAGs.....	12
2.2. Large Language Models (LLMs), Arquitetura Transformer, Pré-treinamento e Alinhamento.....	15
2.3. Do Prompt à Resposta e a Lógica da Inferência Algorítmica.....	19
2.4. Entre a Estatística e o Sentido: Limitações Intrínsecas e Desafios Epistemológicos dos Modelos Generativos.....	23
2.5. Inteligência Artificial Explicável (XAI) e os Desafios na Mitigação da Opacidade Algorítmica.....	27
3. A DECISÃO JUDICIAL.....	31
3.1. Teoria da Integridade do Direito de Ronald Dworkin.....	34
3.2. Princípios Fundamentais do Processo Civil, Conforme Humberto Theodoro Júnior.....	40
4. DA TÉCNICA À CRÍTICA NO HORIZONTE DECISÓRIO.....	46
4.1. As Fronteiras da Máquina e as Limitações das IAGs na Deliberação Jurisdicional.....	49
4.1.1. A Hermenêutica em Ruínas: A Supressão do Sentido pela Estatística.....	52
4.1.2. O Véu e o Juízo: A Opacidade Algorítmica Frente ao Dever de Fundamentar.....	56
4.2. Riscos Ético-Jurídicos da Automação Judicial.....	61
4.3. A Dissonância Entre Decisões Automatizadas e a Ordem Constitucional Brasileira.....	65
5. CONSIDERAÇÕES FINAIS.....	69
REFERÊNCIAS.....	72

1. INTRODUÇÃO

O avanço exponencial das tecnologias de Inteligência Artificial Generativa, em especial aquelas assentadas sobre arquiteturas de *Large Language Models*, tem deflagrado transformações profundas nas formas de produção simbólica, com ênfase na linguagem escrita. No contexto jurídico, a emergência de modelos capazes de redigir textos normativamente estruturados e semanticamente plausíveis acirra o debate acerca da viabilidade de sua aplicação em funções tradicionalmente atribuídas à autoridade judicante. A hipótese de automatização da redação de decisões judiciais, mediante o uso de sistemas algorítmicos treinados em grandes *corpora* textuais, inaugura uma nova fronteira de tensionamento entre técnica e normatividade, estatística e hermenêutica, cálculo e deliberação.

A presente investigação parte do reconhecimento de que a função jurisdicional, em sua essência, transcende a operação formal de silogismos e repousa sobre pressupostos hermenêuticos, valorativos e argumentativos inassimiláveis por sistemas computacionais. Neste horizonte, o estudo se justifica pela urgência de delimitar os contornos ético-jurídicos que balizam a integração de sistemas generativos ao núcleo da atividade decisória, de modo a evitar o risco de desvirtuamento dos princípios constitucionais que regem o devido processo legal, a fundamentação racional das decisões e a dignidade da função jurisdicional. Para tanto, impõe-se uma análise que ultrapasse o fascínio técnico e enfrente, com rigor crítico, as implicações epistêmicas e ético-jurídicas da delegação decisória a sistemas estatísticos opacos.

O objetivo geral do trabalho consiste em examinar criticamente os limites técnicos, epistemológicos e normativos relacionados ao uso de inteligências artificiais generativas na elaboração de decisões judiciais. Nesse sentido, pergunta-se: é possível delegar a atividade decisória a sistemas de Inteligência Artificial Generativa sem comprometer a legitimidade do Poder Judiciário, de modo a atender os requisitos de fundamentação aos moldes da Constituição da República de 1988 e se assegurar as garantias fundamentais, bem como aos princípios dispostos no Código de Processo Civil de 2015?

Como objetivos específicos, pretende-se: (i) descrever os fundamentos operacionais dos LLMs e o funcionamento das IAGs; (ii) analisar os limites cognitivos, epistêmicos e axiológicos desses sistemas frente às exigências do processo jurisdicional; (iii) investigar os riscos de opacidade, vieses e desresponsabilização institucional inerentes à sua aplicação; e (iv) confrontar a lógica algorítmica de predição estatística com a racionalidade hermenêutica

que estrutura o ato de julgar, à luz de teorias como a integridade do Direito, formulada por Ronald Dworkin, e a os princípios universais do processo, segundo descrição doutrinária de Humberto Theodoro Júnior.

A pesquisa desenvolve-se sob uma abordagem de natureza qualitativa, alicerçada na vertente metodológica jurídico-dogmática, com suporte em revisão teórico-bibliográfica. A estratégia argumentativa adotada é de natureza crítico-dedutiva, articulando pressupostos interdisciplinares que integram teoria do Direito, filosofia hermenêutica, linguística, ciência da computação e ética da tecnologia. O arcabouço teórico mobilizado inclui, de um lado, clássicos e contemporâneos da teoria do Direito e processual civil, e, de outro, literatura científica atualizada nos campos da inteligência artificial, *machine learning*, Processamento de Linguagem Natural, *explainable AI*, arquitetura *Transformer* e epistemologia computacional.

A estrutura do trabalho organiza-se em três capítulos, além desta introdução e posteriores considerações finais. O primeiro capítulo apresenta as bases técnico-conceituais da Inteligência Artificial Generativa, com destaque para os paradigmas computacionais que sustentam os LLMs, os mecanismos de inferência, suas limitações epistêmicas e os desafios da explicabilidade. O segundo capítulo examina a natureza da decisão judicial, à luz de sua estrutura argumentativa, seu dever de fundamentação e sua vinculação aos princípios constitucionais do Estado Democrático de Direito. O terceiro capítulo constitui o núcleo crítico do estudo, no qual se examinam, sob chave jurídico-filosófica, os limites da aplicação de IAGs à atividade decisória, confrontando técnica estatística e racionalidade hermenêutica, opacidade algorítmica e dever de motivação, riscos éticos e conformidade principiológica.

Por fim, as considerações finais apresentam aquilo que há de ser ressaltado do estudo geral, de modo a delinear as implicações normativas e institucionais do uso de IAGs na jurisdição e indicar os contornos de um eventual modelo de governança tecnológica compatível com os marcos do Estado Democrático de Direito.

2. A INTELIGÊNCIA ARTIFICIAL GENERATIVA

O desenvolvimento da Inteligência Artificial Generativa (IAG) traduz o avanço das tecnologias baseadas em *machine learning*¹ e, por sua difusão progressiva, tem suscitado debates em múltiplas áreas do conhecimento. Imersas no universo dos sistemas de inteligência artificial, esses modelos generativos tentam simular os processos intelectivos e as faculdades racionais próprias da consciência², através de dispositivos computacionais. As IAGs, a exemplo do ChatGPT, o Gemini e o Llama, conseguem sintetizar novos conteúdos heterogêneos, a partir da identificação de estruturas e padrões compreendidos em grandes volumes de dados. Esses algoritmos³, em linhas gerais, são concebidos para traduzir e gerar textos, imagens, vídeos, códigos ou outras formas de expressão, além de executar tarefas cognitivas comumente associadas à mente humana (Garcia, 2020), sob o égide argumentativo da eficiência e aumento de produtividade.

A aparente subitaneidade com que a Inteligência Artificial Generativa se popularizou mascara uma evolução conceitual e tecnológica de longa data. As fundações filosóficas para a criação de uma "máquina pensante" foram lançadas ainda em meados do século XX, com o precursor teste⁴ proposto por Alan Turing, por meio de seu artigo Computadores e Inteligência⁵, para avaliar a capacidade de uma máquina exibir comportamento inteligente equivalente ao de um ser humano. As primeiras tentativas de simular a conversação, como o

¹ Segundo CETAX: "Machine Learning ou Aprendizado de Máquina é um método de análise de dados que automatiza o desenvolvimento de modelos analíticos, por meio de algoritmos que aprendem iterativamente a partir de dados, permitindo assim que os computadores encontrem insights ocultos sem serem explicitamente programados para procurar algo específico." in CETAX. Machine Learning: o que é, conceito e definição. 6 nov. 2024. Disponível em: <https://cetax.com.br/machine-learning/>. Acesso em: 11 mai, 2025.

² Como lecionam Friedenberg e Silverman (2006, p. 49, tradução livre do original), a consciência pode "ser considerada como nossa percepção subjetiva individual de estados mentais. Esses estados incluem sensação, percepção, imagens visuais, processos de pensamento consciente, emoções e senso de identidade, apenas para citar alguns".

³ Um algoritmo pode ser definido como um procedimento computacional bem estruturado e não ambíguo, que recebe determinados valores como entrada (*input*) e, por meio de uma sequência finita de etapas lógicas e matemáticas, produz valores de saída (*output*). Trata-se, portanto, de um conjunto de regras ou instruções que descrevem, de maneira precisa, o processo de transformação do dado inicial em um resultado final. Em termos gerais, um algoritmo é uma ferramenta para resolver problemas computacionais bem definidos. (CORMEN et al., 2009)

⁴ Alan Turing propôs o "jogo da imitação" em que um juiz, sem saber quem é o humano e quem é o computador, tenta identificar qual dos dois é a máquina, realizando uma série de perguntas por meio de um canal de comunicação neutro; se a máquina conseguir convencer o juiz de que também é humana em uma proporção significativa de interações, ela passa no teste, o que, para Turing, seria um indicativo prático de inteligência artificial.

⁵ TURING, Alan M. Computing machinery and intelligence. Mind, 1950.

programa ELIZA⁶, ainda na década de 1960, operavam com base em regras e reconhecimento de padrões pré-programados, constituindo um simulacro de diálogo sem, contudo, possuir capacidade de aprendizado ou de geração autônoma.

Desde os primeiros experimentos com modelos probabilísticos⁷ para geração de texto até a concepção de Redes Generativas Adversariais (GANs) em 2014 e, mais decisivamente, a criação da arquitetura *Transformer* em 2017, a capacidade de emular a criatividade humana foi paulatinamente ampliada. Essa evolução histórica pavimentou o caminho para os paradigmas tecnológicos disruptivos que hoje definem o campo. Assim, com o intuito de desvelar essa estrutura, a seção a seguir, dedicar-se-á à análise dos seus pilares, especialmente os Grandes Modelos de Linguagem (LLMs) e as redes neurais que lhes dão forma, explicando como a capacidade generativa se efetiva a partir de fundamentos matemáticos e computacionais.

2.1. Fundamentos Conceituais e Paradigmas Tecnológicos das IAGs

A emergência da geração autônoma de conteúdos e a própria materialização da capacidade generativa, longe de constituir mera abstração conceitual, é fruto direto da conjugação de avanços tecnológicos consolidados e de formulações matemáticas complexas. Para isso, no alicerce desses sistemas de Inteligência Artificial Generativa, se encontram as inovações no desenvolvimento em *deep learning*⁸, com especial relevo, no escopo do presente estudo, àqueles que dizem respeito aos *Large Language Models* que empregam as arquiteturas de redes neurais do tipo *Transformer*, utilizados no desenvolvimento dos *chatbots*⁹.

Tipificado por um sistema de equações não lineares com elementos de descontinuidade, o arcabouço matemático tácito dessas arquiteturas acaba por ser capaz de

⁶ Desenvolvida no MIT por Joseph Weizenbaum, ELIZA simulava um psicoterapeuta ao usar simples regras de correspondência de padrões para transformar as falas dos usuários em perguntas reflexivas, incentivando-os a explorar seus sentimentos e demonstrando como a manipulação textual pode criar a impressão de empatia artificial.

⁷ No período entre as décadas de 1980 e 1990, foram as redes neurais recorrentes (RNNs) que ganharam notoriedade por sua capacidade de processar sequências de dados e gerar textos.

⁸ Segundo o Data Science Academy: “Deep Learning usa camadas de neurônios matemáticos para processar dados, compreender a fala humana e reconhecer objetos visualmente. A informação é passada através de cada camada, com a saída da camada anterior fornecendo entrada para a próxima camada. A primeira camada em uma rede é chamada de camada de entrada, enquanto a última é chamada de camada de saída. Todas as camadas entre as duas são referidas como camadas ocultas. Cada camada é tipicamente um algoritmo simples e uniforme contendo um tipo de função de ativação”. in DATA SCIENCE ACADEMY. Capítulo 3 – O que são redes neurais artificiais profundas ou Deep Learning? *Deep Learning Book*, 2025. Disponível em: <https://www.deeplearningbook.com.br/o-que-sao-redes-neurais-artificiais-profundas/>. Acesso em: 10 mai, 2025.

⁹ Chatbots são ferramentas que mimetizam interações entre humanos e assistentes, de modo a responder perguntas e simular conversas interativas com os usuários por meio de linguagem natural.

mapear entradas e saídas de forma a reproduzir os exemplos de seu treinamento. No entanto, em contraste aos mecanismos predecessores de Inteligência Artificial, restritos à classificação ou predição de informações paralelas aos dados de entrada, as IAGs se distinguem em virtude do seu potencial para produzir saídas inéditas, extrapolando os padrões internalizados durante o processamento do *dataset*¹⁰ de treinamento. Com efeito, a toda evidência, é essa a característica fundamental que lhe confere notoriedade: a faculdade de não se limitar à reprodução literal de exemplos, mas, pelo contrário, de consubstanciar novas combinações de informações estruturadas.

As redes neurais¹¹, assim, formam a espinha dorsal dos sistemas de GenAI, operando como simulacros dos processos neuronais do cérebro humano. Sua estrutura auto-organizada, composta por camadas interconectadas de neurônios artificiais, é capaz de identificar fenômenos, ponderar alternativas e, por fim, formular conclusões. Em vista disso, a execução operacional do *software* generativo repousa sobre modelos estatísticos de natureza probabilística, que, treinados com vastos e diversos repositórios de dados (e.g., conteúdo disponível na Internet), podem identificar padrões constitutivos da informação, apreender sequências articuladas, captar estruturas gramaticais regulares e discernir relações contextuais que permeiam a comunicação (Google Research, 2022). Dessa forma, por meio de algoritmos de predição e vetores probabilísticos, o modelo estima o próximo elemento em uma determinada cadeia textual, gerando, com isso, respostas aparentemente coerentes¹², estruturalmente plausíveis e semanticamente verossímeis a comandos ou *prompts* fornecidos por seus usuários.

No plano funcional, o ciclo de produção de conteúdo por uma IAG pode ser decomposto em três fases operacionais basilares: (i) o treinamento inicial (*pre-training*), fase na qual o modelo é exposto a vastos conjuntos de dados multilíngues e códigos-fonte, com vistas à internalização de padrões discursivos e inter-relações semânticas; (ii) o ajuste fino (*fine-tuning*), destinado a otimizar o desempenho em tarefas específicas e à adaptação a

¹⁰ Dataset é o conjunto estruturado de informações, tipicamente disposto em tabelas, matrizes ou formatos especializados, que permite recuperação eficiente e análise sistemática.

¹¹ Segundo IBM: “Cada rede neural consiste em camadas de nós ou neurônios artificiais, uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. Cada nó se conecta a outros e tem seu próprio peso e limite associados. Se a saída de qualquer nó individual estiver acima do valor limite especificado, esse nó é ativado, enviando dados para a próxima camada da rede. Caso contrário, nenhum dado é passado para a próxima camada da rede.” in IBM. What is a neural network?. IBM Think, 2021. Tradução livre do original. Disponível em: <https://www.ibm.com/think/topics/neural-networks?>. Acesso em: 10 mai, 2025.

¹² Bender et al. (2021, tradução livre do original) alertam para a ilusão de coerência nos textos gerados por modelos de linguagem, os quais classificam como apenas aparentemente coerentes. Conforme advertem, “a tendência dos interlocutores humanos de imputar significado onde não há nenhum pode enganar tanto os pesquisadores de PLN quanto o público em geral, fazendo-os tomar o texto sintético como significativo. (...) Dizemos aparentemente coerente porque a coerência está, de fato, nos olhos de quem vê”.

molduras particulares, conferindo à ele uma especialização direcionada; e, por fim, (iii) a inferência, momento em que, com base nos padrões estatísticos apreendidos e armazenados em seus parâmetros sinápticos, o modelo produz novos conteúdos em resposta direta aos comandos ou *prompts* fornecidos pelo seu interlocutor.

O senso comum¹³ frequentemente exalta a capacidade adaptativa dessas tecnologias e, em virtude disso, apontam sua inserção progressiva em múltiplos setores da economia, ciência e da cultura — invariavelmente com o propósito de emular ou corresponder à amplitude e à flexibilidade do pensamento humano¹⁴. Todavia, não obstante sua aparente versatilidade, as IAGs são, em sua maioria, modelos pré-treinados para a execução de tarefas pontuais, de sorte que, embora apresentem certa elasticidade operacional, permanecem adstritas às funções correlatas à sua tarefa primordial, qual seja, a predição autorregressiva do próximo elemento (e.g., palavra, token ou pixel) em uma sequência (IBM, 2024; Franceschelli; Musolesi, 2024). Nesse sentido, mesmo os outputs considerados "novos" não deixam de ser, em última instância, variações sintéticas de padrões previamente assimilados. Como bem observa Arão (2024), “espantamo-nos com a capacidade da inteligência artificial de copiar o estilo de algum artista, mas não nos atentamos para o fato de que isso é o máximo que ela pode fazer: reconhecer padrões e repetir”.

Com o progressivo aperfeiçoamento dos modelos generativos, resta evidente sua capacidade exponencialmente crescente de produzir textos aparentemente coerentes, detalhados e com estilo que — tão somente — mimetiza os atributos de criatividade¹⁵ e da variabilidade linguística característicos da expressão humana. Entretanto, mesmo que suas aplicações venham a se expandir em diversas esferas, o conceito central que delimita e define a IAG subsiste ancorado à ideia de geração estatístico-preditiva de material linguístico. Estes, que embora pragmaticamente inéditos, são o produto do reconhecimento e da recombinação

¹³ Arão (2024) destaca a importância de se afastar “de todo marketing e toda fetichização que envolve a inteligência artificial”, a fim de compreendê-la em sua essência, um método estatístico de reconhecimento de padrões.

¹⁴ Damásio (2012) argumenta que o pensamento humano deriva da interação entre o corpo e o cérebro, em que emoções e sensações corporais, o que ele denomina como “marcadores somáticos”, guiam a elaboração mental, moldam decisões e conferem coerência ao fluxo da consciência. No mesmo sentido, postula que cérebro e o corpo constituem organismo indissociável e os fenômenos mentais que deles derivam só podem ser cabalmente compreendidos no contexto de sua interação com o ambiente que o rodeia. Portanto, a mente integra necessidades, estados emocionais e representações conscientes, expressando-se como esse contínuo unificado de imagens, sentimentos e intenções.

¹⁵ Conforme concluem Franceschelli e Musolesi (2024, tradução livre do original): “embora os LLMs sejam capazes de valor e uma versão fraca de novidade e surpresa, sua natureza autorregressiva interna parece impedi-los de alcançar a criatividade transformacional. (...) Um processo criativo exigiria motivação, pensamento e percepção, propriedades que os LLMs atuais não possuem”.

de padrões extraídos dos dados previamente processados e estruturados pelos respectivos algoritmos de aprendizado.

2.2. *Large Language Models (LLMs)*, Arquitetura *Transformer*, Pré-treinamento e Alinhamento

É no limiar onde a lógica computacional, regida por estruturas algorítmicas rigorosas, corteja a fluidez da semântica humana, que despontam os *Large Language Models*, universalmente reconhecidos pela sigla LLMs. Estes sistemas materializam uma categoria preeminente de artefatos de aprendizado de máquina, concebidos sob a ambição de espelhar o próprio sopro criativo e interpretativo do intelecto humano, arquitetados para processar e gerar linguagem natural¹⁶ de maneira funcionalmente antropomórfica (IBM, 2024; Walter; Zbinden, 2022). A eficácia desses modelos é forjada durante sua exposição inicial à vastos conjuntos de dados textuais que, conjugados à sua escala massiva de parâmetros, lhes confere a capacidade de identificar padrões linguísticos complexos e, subsequentemente, produzir respostas coesas, aparentemente coerentes e contextualmente ajustados.

Inicialmente, a compreensão integral dos LLMs pressupõe a elucidação da arquitetura nomeada *Transformer*, originalmente proposta por Vaswani *et al.* (2017) no artigo seminal "Attention Is All You Need". Cabe salientar que, no presente estado da arte, as LLMs sob exame, bem como a maioria dos modelos de linguagem modernos, se baseiam preponderantemente na supracitada arquitetura de Rede Neural Profunda (*Deep Neural Networks* - DNNs). O entendimento pacífico é de que essa arquitetura revolucionou o campo do Processamento de Linguagem Natural (*Natural Language Processing* - NLP¹⁷) ao permitir o processamento paralelo de sequências textuais, mediante o emprego de mecanismos de autoatenção (*self-attention*), os quais possibilitam que o modelo considere, de forma

¹⁶ Segundo Friedenberg e Silverman (2006, p. 386, tradução livre do original): “a linguagem natural é a personificação da cognição humana e da inteligência humana. É muito evidente que a linguagem natural inclui uma abundância de frases e declarações vagas e indefinidas que correspondem à imprecisão nos conceitos cognitivos subjacentes. Termos como ‘alto’, ‘baixo’, ‘quente’ e ‘bem’ são extremamente difíceis de traduzir em representação de conhecimento, como exigido para os sistemas de raciocínio em discussão. Sem essa precisão, a manipulação simbólica dentro do computador é, no mínimo, sombria. No entanto, sem a riqueza de significado inerente a tais frases, a comunicação humana seria severamente limitada”.

¹⁷ Natural Language Processing representa a confluência fundamental entre os domínios da Linguística e das Ciências da Computação. Seu objetivo precípua reside no desenvolvimento de sistemas computacionais capazes de processar, analisar, compreender e gerar expressões linguísticas tipicamente humanas. Com efeito, diversas aplicações oriundas da NLP encontram-se disseminadas e integradas ao cotidiano da sociedade contemporânea, exemplificadas por ferramentas como o corretor automático, mecanismos de sugestão preditiva de texto e sistemas de detecção e verificação de plágio.

simultânea e holística, o contexto integral de uma sequência de palavras, ponderando a relevância relativa de cada termo na composição do texto.

Um dos fatores distintivos dessa arquitetura e que sustenta sua eficácia, reside, pois, na eliminação da recorrência e na incorporação das codificações posicionais (*positional encodings*), que injetam informações sobre a ordem relativa ou absoluta dos elementos em uma sequência, algo que os mecanismos de autoatenção (*self-attention*), por si sós, não capturam intrinsecamente (Vaswani *et al.* 2017). Este mecanismo de atenção, então, passa a ser responsável pela captação das relações contextuais existentes entre os léxicos de uma frase ou parágrafo. Para isso, cada unidade linguística é representada em vetores em um espaço multidimensional, por meio de representações distribuídas da linguagem conhecidas como *embeddings*¹⁸, os quais delineiam semanticamente as inter-relações entre os termos no espaço vetorial.

Destaca-se que, diferente das arquiteturas predominantes anteriormente, como as Redes Neurais Recorrentes (*Recurrent Neural Networks* - RNNs) e as Redes de Memória de Longo Curto Prazo (*Long Short-Term Memory* - LSTMs), que exibiam limitações naturais à estrutura na modelagem de dependências de longo alcance em sequências textuais e somente processava sequencialmente cada elemento de entrada, a arquitetura *Transformer* prescinde da rígida linearidade e sequencialidade temporal. Através da utilização de mecanismos de atenção distribuída, são capazes de apreender, sem embargo da posição relativa no texto, as complexas interdependências semânticas e sintáticas entre as expressões. Esse aprimoramento técnico culmina em uma percepção significativamente mais profunda das estruturas linguísticas e facilita a modelagem de dependências semânticas de longo prazo, aspecto crucial para o desempenho superior dos LLMs frente a seus antecessores.

No âmbito dos LLMs generativos, a exemplo do GPT-4, é possível observar a prevalência da variante *Decoder-only*¹⁹ da arquitetura *Transformer*. Nesta configuração, a estrutura processa uma sequência de entrada (*input*) e gera, de modo autoregressivo, a saída textual (*output*), prevendo iterativamente o próximo elemento levando em conta os anteriores. Essa capacidade de confecção textual é viabilizada por um processo intensivo de

¹⁸ Segundo Seno et al. (2024): “Os vetores de embeddings podem ser estáticos ou dinâmicos. Os embeddings estáticos permanecem fixos uma vez aprendidos, ou seja, eles não podem ser ajustados ou modificados para uma tarefa específica. Ao contrário desses, os embeddings dinâmicos podem ser ajustados em tarefas específicas, se adaptando às nuances específicas da tarefa e ao contexto atual. A escolha entre essas abordagens depende das necessidades da aplicação, do domínio e das características das tarefas em que os embeddings serão utilizados”.

¹⁹ A arquitetura *Transformer*, originalmente proposta por Vaswani et al. (2017), é estruturada em dois componentes cruciais: o codificador (encoder) e o decodificador (decoder). O módulo codificador é responsável pela análise da integralidade da sequência textual de entrada, a partir da qual constrói uma representação desta. Subsequentemente, o módulo decodificador utiliza a referida representação como insumo fundamental para a geração da sequência de saída correspondente.

pré-treinamento auto-supervisionado (*self-supervised pretraining*), mediante o qual o modelo é exposto a um vasto *corpus* textual, oriundo de diversas fontes e desprovido de rótulos explícitos, o que lhe permite inferir das amostras de entrada os padrões gramaticais, sintáticos e semânticos.

Nessa dinâmica de treinamento, conhecida como *Causal Language Modeling*²⁰, a própria predição subsequente em uma sequência atua como rótulo de treinamento. Isso permite que o modelo aprenda a prever, com acurácia progressiva, a próxima unidade linguística, sem necessidade de intervenção humana direta para rotular os dados. Desse modo, o papel precípua do pré-treinamento, demonstrado anteriormente, é calibrar os valores ótimos para os bilhões de pesos sinápticos da rede, de modo a reproduzir de forma generalizada os exemplos do conjunto de dados. Contudo, como adverte Garcia (2020), a concepção de que o modelo resultante possa abranger a integralidade dos casos de treinamento é recebida com reservas por se considerar que tal modelo pode ter apenas memorizado os exemplos (*overfitting*), sem ter efetivamente extraído o padrão de mapeamento.

Ressalta-se, contudo, que esse aprendizado resulta em uma representação puramente estatística dos padrões linguísticos e do conhecimento²¹ codificado nos dados, sem a capacidade de interpretá-los subjetivamente ou autonomia para expandir suas funções cognitivas ou tarefas executáveis para além daquelas estritamente definidas em seu escopo de treinamento. De fato, conforme a própria OpenAI (2023) esclarece, o GPT-4 — um dos mais avançados modelos generativos atualmente disponíveis — não aprende com a experiência proporcionada pela interação com o usuário, isto é, após concluído o seu pré-treinamento, permanecem congelados e refletem um estado específico do domínio (Franceschelli; Musolesi, 2024). Portanto, não há modificação dos parâmetros internos e é inócua qualquer tentativa de interação para corrigir sua predição, de modo que qualquer ajuste só pode ocorrer por meio de uma nova rodada de treinamento ou *fine-tuning* com novos dados.

Superado o pré-treinamento, os modelos são submetidos a uma fase ulterior de alinhamento, cujo objetivo primário reside em ajustar suas saídas às expectativas humanas de utilidade, precisão e segurança. Esse refinamento emprega duas estratégias principais. A primeira é o Ajuste Fino Supervisionado (*Supervised Fine-Tuning* – SFT), que expõe o

²⁰ SHARDLOW, Matthew; PRZYBYŁA, Piotr. Deanthropomorphising NLP: Can a language model be conscious? PLoS ONE, v. 19, n. 12, 2024. Disponível em: <https://doi.org/10.1371/journal.pone.0307521>. Acesso em: 01 jul. 2025.

²¹ De acordo com o estudo conduzido por Trindade e Oliveira (2024), que buscou avaliar o potencial das inteligências artificiais generativas em contextos de produção acadêmico-científica, os autores concluíram que tais ferramentas “(...) podem ser utilizadas para desenvolver um entendimento básico e inicial sobre um assunto, mas não podem ser utilizadas para compreender o estado do conhecimento científico”.

modelo a conjuntos de dados representativos, criteriosamente selecionados, compostos por pares de *prompts* e respostas alinhadas às variáveis desejadas. A segunda, mais complexa, é o Aprendizado por Reforço com Feedback Humano (*Reinforcement Learning from Human Feedback* – RLHF), em que os avaliadores humanos pontuam as diferentes saídas formuladas como base critérios como clareza e adequação. Dessa forma se estrutura um paradigma de recompensa que orienta o ajuste dos parâmetros internos, reduzindo a divergência entre as previsões do modelo e os dados autênticos. Assim, o procedimento orienta o LLM a gerar saídas que maximizem a pontuação, de modo mitigar deficiências estruturais, como a propensão à elaboração de informações incorretas ou enviesadas, conquanto não elimine por completo as limitações estruturais do dispositivo.

Outro aspecto técnico significativo no funcionamento dos LLMs diz respeito à escala dessas estruturas neurais, que operam com bilhões, ou mesmo trilhões, de parâmetros²² ajustáveis, os quais representam os pesos atribuídos às conexões sinápticas artificiais entre os nós das redes neurais (IBM, 2021; Brown *et al.* 2020). Essa magnitude de componentes ajustáveis, ao passo que permite que os modelos capturem um extenso espectro de relações linguísticas e contextuais, é o que também impõe ao modelo uma obscuridade técnica significativa e elevados custos computacionais para o treinamento e a inferência, o que resulta em desafios à governança da interpretabilidade, ao alinhamento ético e à democratização do acesso pleno a essas tecnologias.

Dentre os exemplos paradigmáticos da atualidade, destacam-se o GPT-4 (*Generative Pre-trained Transformer*), desenvolvido pela OpenAI, notório por sua habilidade em gerar textos longos, aparentemente coerentes e semanticamente adequados; o BERT (*Bidirectional Encoder Representations from Transformers*), criado pelo Google, especializado em tarefas de compreensão das expressões linguísticas, como de classificação de texto, para auxiliar sistemas de busca a interpretar a intenção dos usuários; e o LLaMA (*Large Language Model Meta AI*), da Meta, cujo código aberto foi projetado para rivalizar com seus concorrentes de código fechado e otimizar o desempenho com menor demanda computacional. Esses, entre muitos outros presentes no cotidiano de um número crescente de indivíduos, exemplificam o impacto transformador dos LLMs no progresso das tecnologias de processamento de linguagem natural, constituindo infraestrutura invisível para múltiplas aplicações em diversos setores.

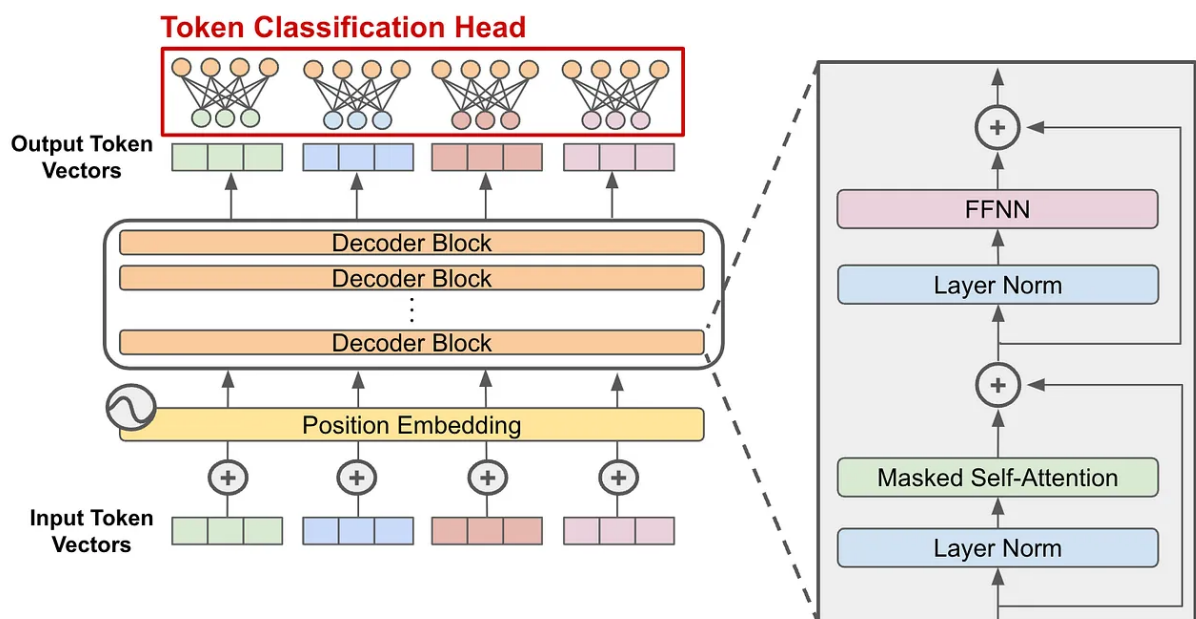
²² O desempenho de um modelo é função crescente da dimensionalidade de seu espaço de parâmetros e do tamanho amostral do conjunto de dados de treinamento. Essa relação demonstra que aumentos na escala de parâmetros e no volume do dataset de treinamento resultam em um aprimoramento significativo das capacidades funcionais do sistema (Google Research, 2022).

2.3. Do Prompt à Resposta e a Lógica da Inferência Algorítmica

A arquitetura basilar dos sistemas de Inteligências Artificiais Generativas, em especial aqueles alicerçados em Grandes Modelos de Linguagem, é edificada a partir das redes neurais profundas, com destaque para a proeminência da arquitetura *Transformers*. Esta arquitetura incorpora uma sofisticada concatenação de camadas interdependentes, cujo condão é transmutar dados linguísticos primevos em representações matemáticas, aptas a codificar as sutilezas semânticas e sintáticas que permeiam os elementos constitutivos da linguagem natural. A partir deste patamar de abstração, podem gerar novas correlações de informações sistematizadas, conforme o arbítrio do operador.

Concluídas as etapas de pré-treinamento e alinhamento das estruturas neurais, inaugura-se a fase de inferência, reputada como o estágio culminante do processo operacional dos sistemas de generativos. Nesta fase, em apertada síntese, o modelo recebe uma sequência textual de entrada que é submetida ao processo de tokenização. Ato contínuo, os tokens resultantes são convertidos em representações numéricas, denominadas *embeddings*. Estes são processados através de múltiplas camadas de autoatenção e redes *feedforward*, permitindo o cálculo autorregressivo das probabilidades dos tokens subsequentes (OpenAI, 2023). Por fim, o sistema produz, de forma iterativa e token a token, a sequência de saída que corresponde ao texto elaborado.

Figura 1 – Estrutura de camadas e operação da arquitetura *Transformer Decoder-Only*.



Fonte: Wolfe (2024)

A *a priori*, o funcionamento de um modelo generativo se deflagra com o recebimento do *prompt*, como uma interrogação, uma instrução ou qualquer formulação textual inicial que estabeleça as condições e norteie a Inteligência Artificial Generativa quanto à natureza da resposta a ser produzida. Em sequência, ocorre a tokenização do texto de entrada, processo que segmenta o conteúdo em unidades mínimas conhecidas como tokens — que correspondem à palavras, subpalavras ou caracteres, conforme as especificações do algoritmo. Esses tokens, por sua vez, são convertidos em *embeddings*, vetores densos que quantificam numericamente cada unidade lexical. Assim, no espaço vetorial resultante, *embeddings* de termos semanticamente próximos tendem a ocupar regiões também próximas, habilitando o modelo a inferir relações de semelhança.

Para modelar as interdependências contextuais em relação às posições relativas e às dependências entre tokens, não apenas ao nível de cada token isolado, são agregados aos *embeddings* os vetores de codificação posicional (*positional encodings*). Tais vetores fornecem informação sobre o índice ordinal de cada token na sequência, permitindo ao modelo distinguir a ordem dos componentes linguísticos²³. A combinação de *embeddings* e codificações posicionais é processada por blocos decodificadores (*Decoder Blocks*), cada qual composto por duas operações principais: o mecanismo de autoatenção (especificamente, a *masked multi-head attention*) e a rede *feedforward* (*Feed Forward Network*).

O mecanismo de autoatenção faz com que cada token, durante seu processamento, pondere a influência de todos os demais na sequência, mediante o cálculo de pontuações de atenção que quantificam a relevância contextual de cada token para a predição do termo ulterior. Assim, um escore de atenção elevado denota uma contribuição proporcional mais significativa do valor associado para a determinação da saída, relevando os focos de importância na cadeia de entrada. Para esse fim, são utilizadas as matrizes conhecidas como *query* (Q), *key* (K) e *value* (V), que, por intermédio de operações matriciais, determinam os pesos de atenção atribuídos a cada token. Em resumo, esse processo consiste em mapear a “consulta” (Q) em relação a um conjunto de “chaves” (K), derivadas do conjunto de treinamento, para depois ponderar os “valores” (V) mais pertinentes à construção da sequência (Vaswani *et al.* 2017).

²³ Na decomposição analítica da abordagem clássica do Processamento de Linguagem Natural, conforme delineado por Barbosa *et al.* (2017), essa etapa corresponderia à denominada Análise Pragmática, por meio da qual o termo é interpretado em consonância com o contexto em que se insere. Trata-se de um desdobramento sucessivo das fases anteriores, iniciadas a partir da tokenização, que culmina em uma cadeia estruturada de procedimentos analíticos, compreendendo a Análise Léxica, Sintática e Semântica.

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

Fonte: Vaswani *et al.* (2017)

A *masked multi-head attention* amplia essa capacidade, replicando múltiplas vezes o cálculo de atenção, de forma paralela, sendo que cada “cabeça” atua com projeções lineares distintas de Q, K e V. Adicionalmente, emprega-se uma máscara²⁴ que obsta a consulta de elementos subsequentes durante a inferência, na medida em que o modelo deve considerar exclusivamente o contexto precedente para a predição do próximo token, assegurando a coerência unidirecional necessária à tarefa de composição. As saídas oriundas de cada cabeça de atenção são, então, concatenadas e submetidas a uma transformação linear adicional, que consolida a saída.

Quanto à *Feed Forward Network* (FFN), esta realiza transformações não lineares de modo independente para cada posição da cadeia. Tipicamente, uma FFN em um bloco *Transformer* consiste em duas camadas lineares, a entrada e a saída, intercaladas por uma camada oculta com uma função de ativação não linear, como a ReLU (*Rectified Linear Unit*) ou, em algumas arquiteturas mais recentes, suas variantes como GELU (*Gaussian Error Linear Unit*). Sua função é processar, sem loops ou ciclos, as representações oriundas da camada de atenção, de modo a extrair e refinar relações semânticas mais complexas antes de repassá-las às camadas subsequentes.

$$\text{FFN}(x) = \max(0, xW_1 + b_1)W_2 + b_2$$

Fonte: Vaswani *et al.* (2017)

Posteriormente às etapas de autoatenção e *feedforward*, são empregadas conexões residuais (*skip connections*) e mecanismos de normalização de camada (*layer normalization*). Essas técnicas objetivam mitigar o problema do desaparecimento de gradiente (*vanishing gradient*²⁵), que pode comprometer a calibragem dos parâmetros em redes profundas, além de conferir estabilidade ao processo por meio da normalização das ativações internas, a fim de evitar que o valor de calibragem tenda a valores extremos.

²⁴ Saiba mais em: <https://blog.dsacademy.com.br/funcionamento-dos-llms-e-machine-unlearning/>. Acesso em: 16 mai. 2025.

²⁵ Saiba mais em: <https://medium.com/@amanatulla1606/vanishing-gradient-problem-in-deep-learning-understanding-intuition-and-solutions-da90ef4ecb54>. Acesso em 28 mai. 2025.

De acordo com Wolfe (2024), cada bloco de processamento — compostos pelo mecanismo de autoatenção, pela FFN, pelas conexões residuais e pela normalização de camada — é empilhado sucessivamente em profundidade. Como anteriormente exposto, predomina, em muitos modelos generativos textuais, a arquitetura do tipo *Decoder-only*, constituída exclusivamente por blocos decodificadores, pois a tarefa primordial é a elaboração sequencial de texto. Nesse esquema, a entrada, após tokenização e conversão em *embeddings* acrescida das codificações posicionais, é processada sequencialmente através desta pilha de blocos decodificadores. Em cada passo, ao tempo da geração, a predição do token mais provável é apensado ao conjunto de entrada existente, e este novo conjunto serve como *input* para a predição do token posterior. Esse ciclo autorregressivo se perpetua até que um critério de parada seja atingido, análogo a um fio narrativo que se desdobra paulatinamente.

Ao mesmo tempo, à medida que os tokens vão sendo concebidos, são convertidos em palavras ou subpalavras, que, justapostas, compõem a resposta final. Para determinar o percurso de geração textual, de modo a torná-lo menos repetitivo e mais criativo, o modelo emprega heurísticas de decodificação como *beam search*²⁶ ou métodos de *sampling*²⁷, cujo objetivo é equilibrar a propensão à produção do conteúdo estatisticamente mais provável com a introdução de termos menos frequentes. Ao término desse processo de decodificação, a resposta aparentemente coerente e estruturada de acordo com a moldura e as diretrizes iniciais é disponibilizada ao interlocutor.

Cumpra asseverar, por derradeiro, que as IAGs não realizam consultas a bancos de dados, ao contrário dos sistemas tradicionais de recuperação de informação. Sua capacidade de resposta emana de padrões estatísticos, inter-relações semânticas e estruturas linguísticas internalizadas durante o extensivo processo de treinamento, as quais estão distribuídas pela arquitetura neural e codificadas nos bilhões de parâmetros — pesos e vieses das conexões neurais — que constituem o modelo (Brown *et al.* 2020). Destarte, o *software* não armazena frases ou respostas específicas; em vez disso, a partir de sua complexa estrutura matemática, generaliza regras e padrões latentes inferidos a partir do *dataset* de treinamento.

Os pesos ajustados, portanto, conferem à arquitetura a capacidade de responder a perguntas e gerar textos sobre tópicos não explicitamente encontrados durante o treinamento, contanto que estes sejam representáveis por combinações dos padrões previamente assimilados. E, de igual modo, desde que estejam contidos dentro da janela de contexto

²⁶ FRANCESCHELLI, Giorgio; MUSOLESI, Mirco. Creative Beam Search: LLM-as-a-Judge for Improving Response Generation. ICCV, 2024. Disponível em: <https://arxiv.org/abs/2405.00099>. Acesso em 05 mai. 2025.

²⁷ HOLTZMAN et al.. The Curious Case of Neural Text Degeneration. In: International Conference on Learning Representations (ICLR), 2020. Disponível em: <https://arxiv.org/abs/1904.09751>. Acesso em 05 mai. 2025.

máxima²⁸ (*maximum context window*) suportada pela arquitetura, que define a amplitude da sequência de tokens pregressos que o modelo pode efetivamente processar para a geração subsequente. Consequentemente, as respostas geradas, por mais articuladas e convincentes que se apresentem, não são o reflexo de uma compreensão profunda²⁹, de inferência lógica ou raciocínio autêntico (CNJ, 2024), mas, antes, a probabilística reverberação de padrões linguísticos abstraídos do *corpus* de treinamento que delineou sua parametrização.

2.4. Entre a Estatística e o Sentido: Limitações Intrínsecas e Desafios Epistemológicos dos Modelos Generativos

Não obstante o conspícuo potencial tecnológico das Inteligências Artificiais Generativas como instrumentos auxiliares de valor significativo em diversas áreas do conhecimento, é imperioso demonstrar o véu de limitações intrínsecas que se sobrepõem ao seu horizonte promissor. Tais limitações derivam tanto de sua constituição técnica, como a escala massiva de seus fatores, quanto dos pressupostos epistemológicos e ontológicos que fundamentam sua concepção.

Em primeiro plano, cumpre destacar a limitação proveniente da sua dependência umbilical ao *corpus* informacional com o qual foi treinado. A vista disso, a GenAI, em sua essência, não engendra conhecimento *ex nihilo*, ao contrário, identifica e reproduz recorrências probabilísticas e associações linguísticas a partir de informações preexistentes,

²⁸ A janela de contexto (*context window*), designa o número máximo de tokens que o modelo específico pode processar e considerar simultaneamente como entrada para uma única inferência. A janela de contexto delimita a "memória" de curto prazo do modelo para a sequência de entrada e, portanto, há delimitado a quantidade de tokens que permanecerá acessível à arquitetura durante o processamento da sequência textual corrente. Quando uma sequência textual ultrapassa esse limite, os tokens mais antigos ou menos relevantes são descartados ou deixam de ser considerados para a geração da resposta. A título ilustrativo, no estudo conduzido por Brown et al. (2020), no qual se analisa o desempenho do modelo GPT-3, a janela de contexto era de 2048 tokens. Posteriormente, conforme descrito por Koubaa (2023), a arquitetura GPT-4 apresenta versões com janelas de contexto significativamente ampliadas, variando entre 8192 a 32768 tokens, a depender da variante do modelo analisada.

²⁹ O filósofo americano John Searle, através do seu artigo "Minds, brains and programs", formulou o Argumento do Quarto Chinês para questionar a noção de compreensão em sistemas computacionais. Nesse experimento mental, um indivíduo, isolado em uma sala, dispõe apenas de um manual de regras sintáticas que orienta a manipulação de símbolos escritos em chinês — idioma que ele desconhece por completo. O participante recebe sentenças em chinês, aplica estritamente as instruções do manual e devolve respostas igualmente em chinês. Apesar da aparente coerência textual das respostas, o operador interno não adquire nenhum entendimento semântico do idioma. Searle conclui, portanto, que, se o agente humano seguindo as regras não obtém compreensão, um sistema de computador que realize as mesmas operações também não compreende, ainda que seu comportamento seja indistinguível do humano (Searle, 1997).

elaboradas por agentes humanos e historicamente condicionadas³⁰. Importa destacar, ainda, que essas ferramentas não possuem a capacidade de realizar análises substantivas mesmo sobre os temas observados no treinamento e, por consequência, não conseguem identificar lacunas, divergências, equívocos ou inconsistências nas informações ou argumentos (Trindade; Oliveira, 2024), o que se traduz, portanto, em uma mera representação do conhecimento extraído.

Em virtude de sua natureza estocástica, estas tecnologias são privadas de uma compreensão³¹ genuína de conceitos abstratos, não discernem relações de causalidade e carecem de consciência contextual, o que torna processos que exigem revisão sistemática, meta-análise ou exegese crítica inexoravelmente dependentes da intervenção humana (Salvagno; Taccone; Gerli, 2023). Destarte, seu funcionamento se circunscreve ao horizonte de previsibilidade estatística, desprovida de intencionalidade e da capacidade para internalizar significados (Bender; Koller, 2020) ou de vislumbrar as sutilezas culturais e axiológicas que permeiam a experiência humana. Portanto, são absolutamente inaptas a gerar novas ideias, tecer raciocínios críticos ou emitir juízos de valor³² que transcendam o cálculo, o que, em última instância, sentencia à crua redução da “experiência ao cálculo, a contingência à probabilidade, o humano ao dado” (Oliveira, 2025).

Ademais, um óbice estrutural adicional reside na opacidade algorítmica — frequentemente descrita como o problema da “caixa-preta” (*black box*) — que se manifesta na dificuldade, ou mesmo na impossibilidade, de se elucidar com precisão os processos internos e de se traçar um rastreamento causal transparente entre os dados de entrada (*inputs*) e as saídas (*outputs*) elaboradas. As estruturas de grande escala funcionam mediante bilhões ou trilhões de parâmetros, ajustados a partir de vastos conjuntos de dados, de tal forma que as decisões por eles produzidas tornam-se virtualmente ininteligíveis, inclusive para seus

³⁰ Conforme elucida o relatório técnico publicado pela OpenAI (2023), o GPT-4, geralmente, não tem conhecimento dos acontecimentos posteriores a setembro de 2021, data em que se encerrou a maior parte do seu pré-treinamento. O mesmo relatório ressalva, entretanto, que os dados de pré-treinamento e pós-treinamento possuem uma pequena quantidade de dados mais recentes, sem, contudo, especificar a natureza ou a cronologia precisa dessas informações.

³¹ Bender e Koller (2020, tradução livre do original) criticam todas as citações, tanto acadêmicas ou midiáticas, que atribuem aos LLMs a capacidade de “entender” ou “compreender” a linguagem natural ou de apreender o seu “significado”, o que, segundo os autores, são “alegações exageradas causadas por uma incompreensão da relação entre forma linguística e significado”. Conforme sua investigação aponta, a aparente capacidade de “raciocinar” desses modelos “às vezes é uma miragem construída com base em artefatos nos dados de treinamento (ou seja, forma, não significado)”. E completam dizendo: “Nossa contribuição é um argumento, com base teórica, de que um sistema exposto apenas à forma em seu treinamento não pode, em princípio, aprender significado”.

³² Segundo Toledo e Pessoa (2023): “(...) o algoritmo não tem a habilidade de avaliar um dado como bom ou mau, justo ou injusto. Seu código simplesmente reproduz a avaliação feita pelo desenvolvedor ao associar determinado dado com certo modelo de resultado entendido (pelo desenvolvedor) como bom, positivo, correto”.

próprios desenvolvedores, e, com maior razão, para seus operadores e usuários finais. Consoante asseverado pela OpenAI (2024):

Ao contrário da maioria das criações humanas, não compreendemos realmente o funcionamento interno das redes neurais. Por exemplo, engenheiros podem projetar, avaliar e consertar carros diretamente com base nas especificações de seus componentes, garantindo segurança e desempenho. No entanto, as redes neurais não são projetadas diretamente; em vez disso, nós projetamos os algoritmos que as treinam. As redes resultantes não são bem compreendidas e não podem ser facilmente decompostas em partes identificáveis (OpenAI, 2024, tradução livre do original).

Outra problemática fundamental reside na reprodução acrítica de vieses indissociáveis dos dados de treinamento. Os sistemas de GenAI, ao serem treinados com o repositório de informações histórica e socialmente condicionados, sujeitos a limitações de atualização ou lacunas de abrangência temporal, no entendimento de Trindade e Oliveira (2024), são propensos a internalizar e perpetuar vieses e distorções discriminatórias prevalentes no tecido social. Tais ferramentas, destituídas de capacidade para o juízo crítico, não apenas espelham essas assimetrias, mas frequentemente as amplificam, reiterando desigualdades estruturais sob a égide de uma aparente e pretensa neutralidade técnica (Garcia, 2020). Assim, mesmo na hipótese de uma curadoria criteriosa de conjuntos de dados virtualmente neutros para o processo de treinamento ou para a formulação de *prompts*, a própria seleção e delimitação desse *corpora* são mediadas pela subjetividade humana. Consequentemente, subsiste o risco de que, submetido à taxonomias antigas e conservadoras, o agente humano, ainda que inadvertidamente, induza o sistema a reproduzir padrões discriminatórios idiossincráticos, refletindo seus próprios vieses latentes³³ (Bragança; Bragança, 2019).

Aos desafios já delineados, soma-se o fenômeno conhecido como "vazamentos semânticos" ou "alucinações", tipificado pela geração de *outputs* que apresentam inconsistências internas, erros factuais ou asserções carentes de fundamento empírico. Nesse cenário, Inteligência Artificial Generativa, ao realizar a predição do token subsequente com base no contexto precedente, pode amalgamar informações, conceitos e arcabouços linguísticos de modo errático, resultando em saídas carentes de coerência lógica ou fidelidade factual. É digno de nota que o *software* não detém qualquer forma de consciência (Shardlow; Przybyla, 2024; Walter; Zbinden, 2022), sobretudo quanto a veracidade ou falsidade do

³³ No mesmo contexto Bragança e Bragança (2019) alertam: “deste modo, fica claro que a constituição dos sistemas se faz a partir das preferências dos programadores ou dos adquirentes do produto e são reflexo das opiniões e prioridades desses sujeitos. A máquina, por conseguinte, não é inteiramente isenta, uma vez que depende de uma atividade humana preliminar de seleção de informações; o que gera uma preocupação quanto a esses desvios cognitivos.”

conteúdo que enuncia³⁴, e o processo de inferência estatística é alheio ao significado, ao propósito teleológico ou à relevância axiológica do que é gerado. Dessa forma, informações incorretas, obsoletas ou inconsistentes, presentes no conjunto de dados iniciais, podem ser incorporadas e reproduzidas nos *outputs*.

Outrossim, quando confrontada com tarefas que divergem significativamente dos padrões assimilados durante o treinamento, a IA pode incorrer em erros, gerar respostas incoerentes ou, inclusive, produzir informações "alucinadas" (OpenAI, 2023). Entre 2023 e 2025, em meio à crescente adoção dessas tecnologias, foram noticiados diversos casos que ilustram os vazamentos semânticos. Por exemplo, a utilização de inteligência artificial por um Juiz Federal do Tribunal Regional Federal da 1ª Região para a elaboração de minuta de ato judicial³⁵, a investigação de Juiz de Direito pela redação de sentenças mediante algoritmo³⁶ e a citação de dispositivo inexistente do Código Civil em Acórdão do Tribunal de Justiça do Estado de Mato Grosso³⁷, inegável, assim, os riscos inerentes à confiança excessiva depositada por seus operadores nas respostas providas por tais sistemas (CNJ, 2024).

No domínio específico da atividade jurisdicional, urge destacar a incapacidade das GenAI de efetuar a operação lógico-hermenêutica da subsunção normativa (CNJ, 2024), a qual representa o cerne estruturante do raciocínio jurídico. A subsunção, enquanto operação primordial para a concretização do Direito, exige a interpretação das normas jurídicas, a contextualização fática e a valoração das especificidades do caso concreto, com o fito de tornar individual a norma jurídica geral e abstrata. As idiosincrasias de cada lide conferem singularidade à situação *sub judice* e ao sujeito, em um processo deliberativo dialético que supera a identificação estatística de padrões e a derivação de *outputs* opacos.

Conforme anteriormente exposto, os modelos de linguagem de grande escala estão distantes do raciocínio humano (Arão, 2024) e carecem, pois, da sensibilidade hermenêutica³⁸ para a compreensão do significado profundo do texto, para preterir condições fáticas e jurídicas singulares do caso concreto (Toledo; Pessoa, 2023). Estão, sobretudo, limitados estruturalmente à janela de contexto, que os impede de fazer a conexão entre um número elevado de informações e documentos simultaneamente, o que os obsta de realizar a operação

³⁴ ZHAO, Yukun et al. Knowing What LLMs Do Not Know: A Simple Yet Effective Self-Detection Method. arXiv:2310.17918v2, 2024. Disponível em: <https://arxiv.org/abs/2310.17918>. Acesso em: 7 jul. 2025.

³⁵ Saiba mais em: <https://www.migalhas.com.br/quentes/396836>. Acesso em: 10 fev. 2025.

³⁶ Saiba mais em: <https://cj.estrategia.com/portal/juiz-robo-magistrado-sentencas-ia>. Acesso em: 15 mai. 2025.

³⁷ Saiba mais em: <https://cj.estrategia.com/portal/acordao-dispositivo-inexistente-inteligencia-artificial>. Acesso em: 15 mai. 2025.

³⁸ Segundo Streck (2014, p. 261): “hermenêutica significa, tradicionalmente, teoria ou arte da interpretação e compreensão de textos, cujo objetivo precípua consiste em descrever como se dá o processo interpretativo-compreensivo”.

axiológica e teleológica que subjaz à atividade judicante. Embora demonstrem capacidade de reproduzir com verossimilhança o léxico técnico-jurídico e as estruturas argumentativas próprias do discurso jurídico, os *outputs* gerados por esses *softwares* são desprovidos de fundamentação racional e reflexão crítica, elementos indissociáveis da função jurisdicional.

A valoração axiológica, por seu turno, configura-se como outro obstáculo intransponível para os sistemas de IAG. A ponderação entre princípios jurídicos não é passível de redução à operações matemáticas ou formulações algorítmicas, esta requer uma apreciação racional da relevância normativa de cada princípio, à luz das circunstâncias do caso concreto, inserida em uma moldura de deliberação humana que engloba escolhas conscientes, conjecturas contrafactuais³⁹ e a consideração de argumentos contrapostos. Em síntese, o desempenho de atividades jurídicas pressupõe a interpretação (Streck, 2020). Os sistemas de IAG, dada a sua natureza algorítmica, operam mediante cálculos estatísticos, sendo, portanto, incapazes de apreender o significado normativo ou de atribuir sentido jurídico aos dados processados. Restringem-se, por conseguinte, à replicação de padrões linguísticos, carecendo da dimensão valorativa e teleológica que é consubstancial ao ato decisório, o que, em última análise, limita o processo de ponderação a um exercício de replicação probabilística.

2.5. Inteligência Artificial Explicável (XAI) e os Desafios na Mitigação da Opacidade Algorítmica

No debate quanto à opacidade algorítmica, característica nativa de arquiteturas com larga escala de parâmetros, emerge a Inteligência Artificial Explicável (*Explainable Artificial Intelligence* – XAI). Esta subárea da Inteligência Artificial se dedica ao desenvolvimento de métodos e técnicas que visam elucidar os processos de decisão e os resultados gerados por sistemas de IAG, com fito de torná-los transparentes, interpretáveis e compreensíveis aos seres humanos (Arrieta *et al.* 2020). A inteligibilidade das predições algorítmicas, nesse sentido, passa a ser condição fundamental para a validação da confiabilidade do sistema, para a detecção e mitigação de vieses, para a atribuição de responsabilidade institucional e, sobretudo, para o fortalecimento da confiança pública e da aceitabilidade social dessas tecnologias.

³⁹ Chomsky, Roberts e Watumull (2023, tradução livre do original) caracterizam a falha mais profunda desses sistemas “é a ausência da capacidade mais crítica de qualquer inteligência: dizer não apenas o que é, o que foi e o que será — isto é, descrição e previsão —, mas também o que não é e o que poderia e não poderia ser. Esses são os ingredientes da explicação, a marca da verdadeira inteligência”.

A adoção de métodos de XAI, portanto, tem como objetivo atender, de forma técnica, às exigências por inteligibilidade e auditabilidade do raciocínio subjacente às previsões algorítmicas (Maranhão, 2019), bem como das conexões semânticas que culminam em determinadas conclusões. Esses recursos permitem examinar a consistência lógica, a coerência estrutural e a legitimidade material das inferências assistidas pelo algoritmo. Paralelamente, a própria densidade estrutural dos LLM e o elevado número de parâmetros que os compõem, compromete sobremaneira a rastreabilidade dos seus processos internos, convertendo a tarefa de reconstrução explicativa em um empreendimento técnico-cognitivo de notória complexidade.

No esforço de suprir a demanda por auditabilidade, a XAI tem se estruturado a partir de diferentes estratégias metodológicas. Entre elas, destacam-se os chamados Métodos Agnósticos ao Modelo (*Model-Agnostic*), concebidos para serem aplicáveis a qualquer tipo de arquitetura, independentemente de sua forma ou natureza interna. Tais abordagens usualmente recorrem à construção de um segundo modelo, mais simples e interpretável, que busca mimetizar o comportamento do sistema original, ainda que de modo aproximativo (Rudin, 2019). Dentre os exemplos de sua execução incluem o LIME⁴⁰ (*Local Interpretable Model-Agnostic Explanations*), que aproxima, em âmbito local, o comportamento de arquiteturas complexas por meio de modelos substitutos simplificados e interpretáveis, e o SHAP⁴¹ (*Shapley Additive Explanations*), que atribui valores de importância (valores de Shapley) a cada variável de entrada, permitindo uma alocação consistente da contribuição de cada característica para as previsões do modelo.

Contudo, é preciso reconhecer que as explicações produzidas por tais métodos constituem, em si mesmas, configurações aproximativas do comportamento do sistema original, o que instiga a reflexão acerca da fidelidade e completude em representar fidedignamente os intrincados processos de inferência da GenAI. Como sustenta Rudin

⁴⁰ LIME é um algoritmo interpretável que pode explicar as previsões individuais de qualquer modelo caixa preta, aproximando localmente seu funcionamento a um modelo interpretável. RIBEIRO, Marco Tulio; SINGH, Sameer; GUESTRIN, Carlos. Why should i trust you?: Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016. p. 1135-144.

⁴¹ Segundo Awan (2023, tradução livre do original): O SHAP explica a previsão de uma amostra de dados calculando a contribuição de cada recurso para a previsão do algoritmo. “No aprendizado de máquina, cada atributo recebe um valor de importância que representa sua contribuição para o resultado do modelo”. in AWAN, Abid A. An Introduction to SHAP Values and Machine Learning Interpretability. Datacamp, 2023. Disponível em: <https://www.datacamp.com/tutorial/introduction-to-shap-values-machine-learning-interpretability>. Acesso em: 07 jul, 2025.

(2019), modelos que exigem explicações externas para tornarem-se compreensíveis talvez deveriam, desde sua concepção, ser projetados de forma intrinsecamente interpretável.

Em paralelo, destacam-se os Métodos Específicos ao Modelo (*Model-Specific*), cuja aplicabilidade se restringe a determinadas arquiteturas. No caso das arquiteturas particulares de LLMs, a análise dos mecanismos de autoatenção permite visualizar os pesos atribuídos a diferentes segmentos textuais durante a geração de uma predição, oferecendo, assim, uma espécie de cartografia dos elementos mais influentes no processo inferencial. Além disso, técnicas *post-hoc* de interpretabilidade têm sido utilizadas para representar, a posteriori, o comportamento do modelo por meio de visualizações, análises das representações internas ou estudo das saídas intermediárias. Um exemplo representativo é o uso de estruturas como a Cadeia de Pensamento⁴² (*Chain of Thought*), que busca explicitar os passos intermediários da geração textual para oferecer maior transparência argumentativa.

Não obstante essas iniciativas, permanece o paradoxo central da XAI: a simplificação necessária para tornar uma explicação humanamente inteligível pode mascarar a real complexidade dos mecanismos subjacentes ou, em cenários mais críticos, ser instrumentalizada para conferir um falso verniz de legitimidade a decisões potencialmente enviesadas ou incorretas. Tal tensão é agravada pela diversidade dos interlocutores possíveis — desenvolvedores, magistrados, pesquisadores, operadores do Direito —, cujos níveis de conhecimento técnico e cujas demandas de granularidade variam substancialmente. Como advertem Doshi-Velez e Kim (2017), a utilidade de uma explicação é inseparável do contexto e das competências epistêmicas de quem a recebe, sendo improvável que modelos generalistas de explicabilidade consigam atender simultaneamente a todas essas expectativas.

Mesmo que os defensores da XAI sustentem, com razão, que a inteligibilidade possui o potencial de mitigar os riscos nativos da opacidade, são muitos os desafios a serem superados. Nesse sentido, cumpre ressaltar ainda a tendência geral de inversa proporcionalidade entre desempenho e transparência que assola o desenvolvimento da GenAI (Atrey *et al.* 2025). Modelos mais simples e, portanto, mais interpretáveis tendem a exibir um desempenho inferior em termos de acurácia e capacidade de generalização. Em contrapartida, as estruturas neurais de maior complexidade, que apresentam uma performance superior, veem sua transparência ser drasticamente reduzida, retornando ao problema original da

⁴² Segundo o Google Research (2022, tradução livre do original): “o modelo é incentivado a ‘mostrar seu trabalho’ na resolução de novos problemas (semelhante a como seu professor de matemática da quarta série o incentivava a mostrar as etapas envolvidas na resolução de um problema, em vez de apenas escrever a resposta que você encontrou), ajuda os modelos de linguagem a seguir uma cadeia lógica de pensamento e a gerar respostas mais estruturadas, organizadas e precisas”.

opacidade algorítmica. Portanto, a GenAI verdadeiramente confiável está para além do desenvolvimento de mecanismos capazes de explicá-la, impõe-se a construção de um arcabouço ético-operacional capaz de mediar continuamente a relação sinalagmática entre eficácia preditiva e clareza interpretativa.

3. A DECISÃO JUDICIAL

Em sua acepção mais elementar, a decisão judicial é o pronunciamento pelo qual o órgão jurisdicional resolve uma questão posta à sua apreciação. Contudo, sua profundidade transcende a simples solução de controvérsias. Ela é, antes de tudo, um ato de poder do Estado que, ao proferir um comando, impõe uma determinada conduta ou reconhece uma situação jurídica. Sob a égide do Estado Democrático de Direito, seu múnus precípua é conferir segurança jurídica, eficácia e continuidade à ordem jurídica, transformando um conflito de interesses em uma situação de direito consolidado, passível, se necessário for, de execução forçada.

Adicionalmente à sua tarefa de composição de litígios, a prestação jurisdicional irradia efeitos para toda a coletividade, servindo como indicativo da exegese e aplicação do direito, reforçando, por conseguinte, a confiabilidade no primado da lei e o sentimento de justiça. Nesse paradigma, como assinala Menelick de Carvalho Netto (2004, p. 38), exige-se do Judiciário a prolação de decisões que “satisfaçam, a um só tempo, a exigência de dar curso e reforçar a crença tanto na legalidade, entendida como segurança jurídica, como certeza do Direito, quanto ao sentimento de justiça realizada, que deflui da adequabilidade da decisão às particularidades do caso concreto”. A decisão judicial, portanto, desempenha um papel fundamental na própria construção e evolução do sistema jurídico, na medida em que, ao interpretar e aplicar as normas, os órgãos judicantes contribuem para o desenvolvimento da jurisprudência e conformação do Direito às novas configurações sociais.

Apesar da pluralidade de construções teóricas acerca do ato decisório, a essência do provimento jurisdicional ainda repousa na aplicação do direito material à moldura fática descrita nos autos. Tal operação pressupõe, antes de tudo, a adoção de um método hermenêutico de ajustamento entre normas e fatos⁴³, apto a desvelar o sentido normativo do ordenamento, superando a simples literalidade das normas, para então orientá-lo ao paradigma fático delineado. Nesse arcabouço, a decisão judicial, ao acolher ou rejeitar a pretensão deduzida, efetivamente compõe a lide, substituindo a vontade das partes pela vontade da lei declarada no caso concreto e, assim, eleva a generalidade da norma abstrata ao comando singular do caso concreto.

⁴³ Segundo Coelho (2011, p. 7): “É somente graças ao trabalho hermenêutico de ajustamento entre normas e fatos — tarefa em que se fundem, necessária e inseparavelmente, a compreensão, a interpretação e a aplicação dos modelos jurídicos — que se põe em movimento o processo de ordenação jurídico-normativa da vida social, porque é precisamente no ato e no momento da interpretação-aplicação que o juiz desempenha o papel de agente redutor da distância entre a generalidade da norma e a singularidade do caso concreto”.

Contudo, conforme raciocínio que irá se repetir ao longo deste estudo, já advertia Calamandrei (1960) que o magistrado não pode apenas se limitar a ler e aplicar uma solução aprioristicamente contida no diploma legal, porque a ele cabe “buscar em seu íntimo sentido de justiça uma solução ‘caso por caso’, fabricada, por assim dizer, não em série, mas sob medida”⁴⁴. E sobre o papel do juiz completa:

Podemos, portanto, concluir que reduzir a função do juiz a uma simples atividade de aplicar silogismos significa empobrecê-la, torná-la estéril, dissecá-la. A justiça é algo melhor: é a criação que emana de uma consciência viva, sensível, vigilante, humana. É precisamente este calor vital, este sentido de contínua conquista, de vigilante responsabilidade que é necessário apreciar e incrementar no juiz. (Calamandrei, 1960, p. 80, tradução livre do original)

Estruturalmente, o ato de decidir decompõe-se em três etapas indissociáveis e interdependentes: (i) o exame do acervo fático-probatório; (ii) a interpretação e aplicação do direito material⁴⁵, cuja perspectiva teórica, sob o prisma de Ronald Dworkin, será objeto de análise em seção própria; e (iii) a formulação de um comando decisório final. No plano fático, cabe ao juiz realizar o juízo de admissibilidade e de mérito das provas carreadas aos autos, selecionando aquelas pertinentes à controvérsia e que se coadunam com o ônus probatório distribuído e ao pedido formulado pelas partes. Em seguida, da interpretação se extrai da norma o sentido que melhor se harmonize com o arcabouço constitucional e os princípios de um Estado Democrático de Direito, delimitando, por via de consequência, o campo de incidência da norma. Por fim, se promove o cotejo entre os fatos provados e a norma legal tida por aplicável.

Por fim, a formulação do dispositivo sentencial, por sua vez, está inexoravelmente vinculada ao dever de fundamentação, alçado à condição de garantia constitucional pelo artigo 93º, inciso IX, da Constituição da República, que estabelece que as decisões jurisdicionais devem ser obrigatoriamente fundamentadas (Brasil, 1998). Para além da previsão constitucional, o legislador inseriu no núcleo principiológico da legislação infraconstitucional, nos termos do artigo 11º, e o detalhou na estrutura do ato decisório, conforme o artigo 489º do Código de Processo Civil de 2015. Dispositivos que impõem ao magistrado a obrigação de expor, de maneira clara e precisa, as razões de fato e de direito que

⁴⁴ CALAMANDREI, Piero. *Proceso y Democracia*. Tradução de Héctor Fix-Zamudio. Buenos Aires, Ed. Juridicas Europa-america, 1960. p. 64. Tradução livre do original.

⁴⁵ Maximiliano e Marcaro (2022, p. 24) definem que “a Aplicação não prescinde da Hermenêutica: a primeira pressupõe a segunda, como a medicação a diagnose. (...) Aquela é um meio para atingir a esta; é um momento da atividade do aplicador do Direito. Pode a última ser o estudo preferido do teórico; a primeira, a Aplicação, revela o adaptador da doutrina à prática, da ciência à realidade: o verdadeiro jurisconsulto”.

formaram o seu convencimento, exigindo a demonstração da correlação analítica entre o acervo probatório e a norma incidente, o enfrentamento de todas as teses suscitadas pelas partes e a superação dos argumentos contrários (Brasil, 2015), de modo a revelar o substrato valorativo e probatório que sustentam o seu juízo de certeza.

Além do cumprimento de requisitos de ordem formal, a fundamentação ostenta um caráter eminentemente substantivo, devendo apresentar um encadeamento lógico-argumentativo que evidencie a coerência interna da decisão. Uma decisão desprovida de fundamentação não se presta à crítica e ao controle pelas partes e instâncias superiores, comprometendo a prestação jurisdicional e ensejando, em última análise, a sua nulidade. Ainda, no plano teleológico, o magistrado deve atentar para os fins sociais e as exigências do bem comum a que a norma se dirige, conforme preconiza o artigo 5º da Lei de Introdução às Normas do Direito Brasileiro (Brasil, 2013), alinhando o provimento ao modelo de sociedade que o ordenamento jurídico objetiva concretizar.

Ressalte-se, contudo, que não obstante o esforço de matriz pragmática em decompor a lógica por trás da decisão judicial, como sendo etapas metodológicas para o magistrado, não haveria de se separar do indivíduo por trás da decisão tudo aquilo que o faz ser quem é. Nesse sentido, à luz de Carlos Alberto Menezes Direito (2001, p. 135):

O que se quer significar com isso é que a decisão judicial é, essencialmente, uma decisão humana. Sendo uma decisão humana ela não está, por inteiro, no domínio da ciência ou da técnica. O homem não existe somente porque tem o suposto domínio da razão. O homem existe porque ele é razão e emoções, sentimentos, crenças. A decisão judicial é, portanto, uma decisão que está subordinada aos sentimentos, emoções, crenças da pessoa humana investida do poder jurisdicional. E a independência do Juiz está, exatamente, na sua capacidade de julgar com esses elementos que participam da sua natureza racional, livre e social. (Direito, 2001, p. 135)

Com efeito, a decisão judicial, ao articular de modo estruturado a hermenêutica normativa constitucionalmente adequada, a ponderação de princípios e o rigor probatório, constitui a própria condição de legitimidade do ato. É, pois, no dever de fundamentação que o poder do julgador encontra seu limite e sua justificação, assegurando que o pronunciamento se subsuma a um ato de submissão ao ordenamento jurídico e seus princípios, bem como sirva de instrumento efetivo para a salvaguarda dos direitos fundamentais, afastando qualquer concepção de discricionário exercício de vontade.

3.1. Teoria da Integridade do Direito de Ronald Dworkin

No campo da filosofia jurídica, poucos pensadores provocaram um debate tão intenso sobre o papel do juiz quanto Ronald Dworkin. Sua obra se ergue como uma poderosa crítica não apenas ao positivismo jurídico⁴⁶, mas também ao utilitarismo, que defende a ideia de que o direito deve sempre visar ao maior bem para o maior número de pessoas. Para o presente estudo, importa o edifício teórico construído por Dworkin que posiciona a decisão judicial não como um ato de descoberta mecânica de regras pré-existentes, tampouco como um exercício de pura discricionariedade legislativa, mas como uma atividade eminentemente interpretativa e construtiva.

A teoria da decisão jurídica em Dworkin é indissociável de sua concepção mais ampla sobre o que é o próprio Direito. Para ele, o Direito não se resume ao sistema jurídico previsto em lei, conforme proposto pelo positivismo e representado, nas críticas de Dworkin, por Hart, segundo o qual o ordenamento jurídico seria composto por um conjunto de normas explícitas, identificáveis por um teste de *pedigree*⁴⁷, ou um “modelo de e para um sistema de regras”⁴⁸. Dworkin demonstra a insuficiência desse modelo ao introduzir a distinção fundamental entre regras e princípios⁴⁹. As regras, argumenta, operam em uma lógica de “tudo-ou-nada”⁵⁰: se uma regra é válida e as condições para sua aplicação estão presentes, a consequência jurídica que ela determina deve ser aceita; se não, ela em nada contribui para a decisão. Os princípios, por outro lado, não determinam um resultado de forma conclusiva. Eles possuem uma dimensão de peso ou importância⁵¹ e enunciam uma razão que aponta para uma direção adequada ao caso concreto, mas que pode ser superada em determinadas hipóteses, de modo a não se aplicar um princípio sem excluí-lo do sistema jurídico.

⁴⁶ DWORKIN, Ronald. Levando os direitos a sério. Trad. Nelson Boeira. São Paulo: Martins Fontes, 2002. p. XI.

⁴⁷ O pedigree se refere ao método de identificação e validação de regras com base na forma com que foram inseridas no ordenamento jurídico e de como ganharam autoridade.

⁴⁸ DWORKIN, Ronald. Levando os direitos a sério, p. 36.

⁴⁹ Segundo Dworkin (2002, p. 36): “Denomino “princípio” um padrão que deve ser observado, não porque vá promover ou assegurar uma situação econômica, política ou social considerada desejável, mas porque é uma exigência de justiça ou equidade ou alguma outra dimensão da moralidade”.

⁵⁰ DWORKIN, Ronald. Levando os direitos a sério, p. 39.

⁵¹ DWORKIN, Ronald. Levando os direitos a sério, p. 42.

A presença de princípios no Direito é, para Dworkin, o que refuta a tese positivista da discricionariedade nos chamados "casos difíceis" (*hardcases*) — aqueles em que as regras são vagas, ambíguas ou inexistentes. Para o positivismo, nessas situações, o juiz, ao esgotar o direito existente, faria uso de seu poder discricionário e atuaria como um legislador (Dworkin, 2002, p. 49-50), criando direito retroativamente. Dworkin rechaça veementemente essa visão. Para ele, mesmo nos casos difíceis, o juiz não tem a liberdade de legislar⁵²; ele está obrigado a decidir com base nos princípios que informam a prática jurídica da comunidade.

A decisão jurídica, portanto, não é um ato de criação, mas de descoberta dos direitos já existentes e devidos às partes⁵³, por meio da interpretação construtiva da prática institucional. Resta evidente nesse sentido, que a fundamentação de uma decisão jurídica reside precisamente na capacidade do juiz de articular a mais coerente justificação para sua decisão a partir da totalidade do material jurídico disponível — leis, precedentes e, crucialmente, os princípios que os justificam —, assim, interpretação e aplicação são incindíveis. Dworkin sustenta que mesmo nesses cenários o juiz está vinculado pelo direito e tem o dever de encontrar a resposta correta. Essa resposta, contudo, não se encontra na superfície das regras, mas na dimensão mais profunda do ordenamento jurídico, composta por argumentos de princípios⁵⁴.

No mesmo sentido, o Direito deve ser compreendido como um conceito interpretativo, por excelência, e não meramente semântico⁵⁵. É o que propõe Dworkin ao teorizar a armadilha do “agulhão semântico”, em razão da qual se acredita que todos os operadores do direito partilham da mesma concepção sobre o que ele é e o que ele determina, e que apenas a partir dessa unicidade conceitual é que se pode discutir, sensatamente, sobre como aplicá-lo. Na realidade, é justamente em virtude da discordância que existe entre os juristas, sobre uma mesma norma ou conceito, que se busca resolver dialeticamente as divergências sobre o valor e o propósito que lhe dão a devida sustentação. Como Dworkin lembra, no “direito, porém, grande parte das divergências é teórica, não empírica”⁵⁶, o que leva a crer, em razão disso, é

⁵² DWORKIN, Ronald. Levando os direitos a sério, p. 127.

⁵³ DWORKIN, Ronald. Levando os direitos a sério, p. 430.

⁵⁴ DWORKIN, Ronald. Levando os direitos a sério, p. 556.

⁵⁵ Como argumenta Dworkin (2014, p. 41): “as teorias semânticas pressupõem que os advogados e juízes usam basicamente os mesmos critérios (embora estes sejam ocultos e passem despercebidos) para decidir quando as proposições jurídicas são falsas ou verdadeiras; elas pressupõem que os advogados realmente estejam de acordo quanto aos fundamentos do direito”. E que “(...) sustentam o ponto de vista do direito como simples questão de fato, aquele segundo o qual a verdadeira divergência sobre a natureza do direito deve ser uma divergência empírica sobre a história das instituições jurídicas.”

⁵⁶ DWORKIN, Ronald. O império do direito. Trad. Jefferson Luiz Camargo. 3. ed. São Paulo: Martins Fontes, 2014. p. 56.

que os operadores olham para as mesmas normas e, à luz pela qual a interpretam, vêem coisas diferentes.

Sob essa ótica, as divergências quanto a resolução de casos difíceis estão para além delimitar os contornos significantes dos vocábulos em uma norma ou precedente, com base em critérios semânticos que partilham entre si, uma vez que a composição dos litígios não se esgota na exegese linguística. Com efeito, a dialética processual, enquanto instrumento direcionado a influir no édito decisório, supera o plano semântico para se situar no âmbito da hermenêutica construtiva de estabelecimento dos fundamentos para a decisão. Logo, o debate travado pelos sujeitos processuais visa fixar um propósito ou valor aos objetos e práticas sociais subjacentes à norma, “a fim de torná-lo o melhor exemplo possível da forma ou gênero aos quais se imagina que pertençam” (Dworkin, 2014, p. 64).

É essa dinâmica hermenêutica que permite a justaposição entre paradigmas jurídicos e a interpretação construtiva. Já que os paradigmas variam no tempo, conforme a visão de uma comunidade, cabe à interpretação se ajustar a eles, ao passo que os argumentos contrários servirão para demonstrar que ela deixou de incluir ou explicar um caso paradigmático. Nesse sentido, em “decorrência desse papel especial, a relação entre a instituição e os paradigmas da época será estreita a ponto de estabelecer um novo tipo de atributo conceitual” (Dworkin, 2014, p. 89). Ainda assim, mesmo os paradigmas são sujeitos às interpretações que considere melhores outros paradigmas e o deixe de lado.

De modo a refinar a interpretação construtiva como método de análise do Direito enquanto prática social, Dworkin propõe uma decomposição analítica do processo interpretativo em três fases. Primeiro, o intérprete promove a etapa pré-interpretativa, em que “são identificados as regras e os padrões que se consideram fornecer o conteúdo experimental da prática”. Posteriormente, há de se realizar a etapa interpretativa que exige que “o intérprete se concentre numa justificativa geral para os principais elementos da prática identificada na etapa pré-interpretativa”. E, finalmente, a etapa pós-interpretativa “ou reformuladora à qual ele ajuste sua ideia daquilo que a prática ‘realmente’ requer para melhor servir à justificativa que ele aceita na etapa interpretativa” (Dworkin, 2014, p. 81-82).

É em razão de sua discordância com as teorias que olham apenas para o passado, como o convencionalismo, ou apenas para o futuro, como o pragmatismo, que Dworkin desenvolve sua ideia mais poderosa: o Direito como Integridade. O convencionalismo preza pelas decisões do passado e, em sua ausência, concede ao juiz a discricionariedade para que se decida casos difíceis que não tenham base em precedente, em convenção anterior⁵⁷. O

⁵⁷ DWORKIN, Ronald. O império do direito, p. 142-143.

pragmatismo⁵⁸, por sua vez, mira o progresso futuro da comunidade de modo a maximizar o bem-estar coletivo, mas ao custo de ignorar a coerência com o passado e os direitos individuais, tornando-se um "direito sem direitos"⁵⁹. A integridade exige que o Direito seja visto como um todo coerente⁶⁰ e a decisão correta é aquela que consta ou deriva dos princípios de justiça, equidade e devido processo legal que oferecem a melhor interpretação construtiva da prática jurídica da comunidade⁶¹.

Portanto, a fundamentação de um veredito judicial, fruto das conclusões pós-interpretativas, na concepção dworkiniana, extrapola a aplicação mecânica de uma norma a um fato. Ela se constitui como um ato de interpretação construtiva e auto-reflexiva⁶². Dworkin propõe a célebre metáfora do "romance em cadeia", em comparação entre o direito e a literatura, para ilustrar a tarefa do juiz. Cada juiz, ao decidir um caso, é como um autor convidado a escrever um novo capítulo de um romance que já foi iniciado por outros. Ele não tem a liberdade de escrever o que bem entender, pois, mesmo que supere aquilo que seus antecessores redigiram, seu capítulo deve respeitá-los, ser coerente com a interpretação construída e, ao mesmo tempo, deve fazer do romance a melhor obra literária possível⁶³.

Traduzindo a metáfora do "romance em cadeia", a decisão judicial deve satisfazer duas dimensões simultaneamente: a dimensão da adequação e a dimensão da justificação (Dworkin, 2014, p. 286). A dimensão da adequação exige que a interpretação do juiz seja consistente com a história institucional e com os materiais jurídicos existentes. Ele não pode ignorar o texto da lei ou os precedentes firmados. A dimensão da justificação, por sua vez, impõe que, dentre as várias interpretações que se ajustam minimamente aos dados do sistema, o juiz escolha aquela que apresenta o direito da forma moralmente mais atraente, ou seja, que o justifique à luz dos princípios de justiça, equidade e devido processo legal que informam a comunidade política. Como observa Carvalho Netto (2004, p. 39), "A sensibilidade do juiz para as especificidades do caso concreto que tem diante de si é fundamental, portanto, para que possa encontrar a norma adequada a produzir justiça naquela situação específica".

Esta concepção aperfeiçoa a noção de subsunção. Para Dworkin, portanto, a subsunção não é um processo silogístico simples de operação lógico-dedutiva e mecânica,

⁵⁸ DWORKIN, Ronald. O império do direito, p. 185.

⁵⁹ DWORKIN, Ronald. O império do direito, p. 195.

⁶⁰ DWORKIN, Ronald. O império do direito, p. 203.

⁶¹ DWORKIN, Ronald. O império do direito, p. 272.

⁶² Como observa Streck (2014, p. 364), "o intérprete do Direito é um sujeito inserido/jogado, de forma inexorável, em um (meio) ambiente cultural-histórico, é dizer, em uma tradição. Quem interpreta é sempre sujeito histórico concreto, mergulhado na tradição. Para ter acesso a um texto (e compreendê-lo), é impossível ao intérprete fazê-lo como se fosse uma mônada psíquica, utilizando o cogito herdado da filosofia da consciência. O intérprete é já, desde sempre, integrante do mundo linguístico".

⁶³ DWORKIN, Ronald. O império do direito, p. 276.

pois a própria identificação da premissa maior — a norma aplicável — já é o resultado de um complexo e contínuo exercício interpretativo. Não se trata de pegar uma regra pronta na prateleira do ordenamento e enquadrá-la tecnicamente a um fato. Trata-se de, a partir de todos os fatos do caso, perscrutar continuamente toda a teia do direito, mesmo aquele material que o próprio magistrado afirme ter interpretado com sucesso, para construir a solução que melhor expresse o direito como integridade⁶⁴. A racionalidade jurídica, nesse sentido, não é uma racionalidade lógico-formal, mas sim uma integridade interpretativa e argumentativa de modo que, segundo Dworkin (2014, p. 265), “incentiva um juiz a ser mais abrangente e imaginativo em busca de coerência com o princípio fundamental”.

A racionalidade de uma decisão judicial, portanto, não se mede por sua eficiência em alcançar um objetivo social, mas por sua coerência principiológica com a estrutura de direitos e deveres estabelecida pela história institucional da comunidade, sua reconstrução e aplicação, a toda evidência, a partir da particularidade do caso concreto. É desse modo que Dworkin (2014, p. 292) propõe que a “interpretação tem por finalidade mostrar o que é interpretado em sua melhor luz possível, e uma interpretação de qualquer parte de nosso direito deve, portanto, levar em consideração não somente a substância das decisões tomadas por autoridades anteriores, mas também o modo como essas decisões foram tomadas: por quais autoridades e em que circunstâncias”.

Para ilustrar o método dessa racionalidade ideal, Dworkin cria a figura do juiz Hércules, jurista de capacidades sobre-humanas e com tempo infinito a seu dispor, capaz de reconstruir a teoria que melhor justifica todo o direito de sua comunidade e, a partir dela, encontrar a única resposta correta⁶⁵ para cada litígio. Hércules, ao passo da sua capacidade de examinar exaustivamente todos os elementos que compõem o sistema jurídico, os articula de maneira coerente, ajustando os juízos interpretativos à trajetória institucional, sem estar obrigado a seguir precedentes que se mostrarem viciados, e o adequa ao caso concreto. Embora Hércules seja um ideal inatingível⁶⁶, ele serve como modelo metodológico, um ideal

⁶⁴ No entendimento de Dworkin (2014, p. 273), o direito como integridade “é tanto o produto da interpretação abrangente da prática jurídica quanto sua fonte de inspiração. O programa que apresenta aos juizes que decidem casos difíceis é essencialmente, não apenas contingentemente, interpretativo (...). Oferece-se como a continuidade — e como origem — das interpretações mais detalhadas que recomenda”.

⁶⁵ Na visão de Nunes, Pedron e Bahia (2016, p. 193): “(...) Percebe-se aqui que o espaço para a discricionariedade é eliminado para dar lugar a um espaço hermenêutico e argumentativo. Hércules deve convencer a sociedade que confiou aquele caso ao seu julgamento que fez o melhor que podia — sua decisão tem a pretensão contrafática de ser a única resposta adequada ao caso não apenas pelo esforço hermenêutico mencionado, mas também porque o juiz deverá tomar o caso dentro das particularidades do mesmo e não com um “standard” e um tema. Isso não tem como ser feito sem que estejamos dispostos a fundamentar adequadamente a decisão”.

⁶⁶ DWORKIN, Ronald. O império do direito, p. 294.

de postura auto-reflexiva e crítica, que indica a direção e a natureza do raciocínio jurídico: um esforço hercúleo de interpretação para assegurar que a decisão seja uma emanção da integridade do Direito, e não do arbítrio judicial⁶⁷.

Esta metodologia leva Dworkin a defender a, embora alvo de críticas, tese da única resposta correta. Mesmo nos casos mais intrincados, onde juristas razoáveis divergem, Dworkin sustenta que existe uma única decisão que melhor satisfaz as exigências de ajuste e justificação. O fato de juízes reais, com suas limitações humanas, discordarem ou não conseguirem identificá-la não invalida a sua existência como um ideal regulador. A tese da única resposta correta postula que a mobilização de princípios jurídicos permite ao intérprete conciliar a segurança jurídica com a racionalidade decisória, superando, assim, a aplicação de leis e precedentes que se restrinjam aos seus contornos semânticos. Deste modo, a atividade hermenêutica passa a ser justificada por sua conformidade com os melhores princípios de justiça, equidade e devido processo legal, o que resulta na integridade do direito⁶⁸.

Para Dworkin, recusar a existência de uma resposta correta equivale a subordinar os direitos à aleatoriedade ou ao arbítrio do juiz⁶⁹, uma conclusão que sua teoria rejeita frontalmente. Exige-se, portanto, levar os direitos a sério, reconstruindo os princípios da comunidade em sua melhor luz e respeitando as particularidades de cada caso concreto. Essa postura afasta a ideia de um juiz monologista ou desvinculado da realidade, impondo, ao contrário, um engajamento profundo com a história institucional e com os fatos, que são únicos e irrepetíveis, da lide em questão. Mesmo nos *hard cases*, o autor sustenta que sempre há uma parte que detém o direito de prevalecer, e cabe ao julgador descobrir essa resposta, não criá-la conforme suas preferências. Destarte, o magistrado deixa de ser o simples aplicador de regras ou legislador ocasional, se transformando em intérprete de uma tradição contínua, cujo dever é preservar e aprimorar a integridade do direito, garantindo decisões

⁶⁷ DWORKIN, Ronald. O império do direito, p. 305.

⁶⁸ DWORKIN, Ronald. O império do direito, p. 291.

⁶⁹ Segundo Streck (2013b, p. 359): “(...) o direito enquanto um sistema de regras e princípios não abriria a possibilidade para um juízo discricionário, já que teria sempre uma história institucional a ser reconstruída e que indicaria a melhor decisão a ser tomada”.

fundadas não na vontade do poder⁷⁰, mas na melhor concepção de justiça que a história de sua comunidade pode oferecer.

3.2. Princípios Fundamentais do Processo Civil, Conforme Humberto Theodoro Júnior

A codificação ordinária processual civil brasileira de 2015 representa um divisor de águas na trajetória do direito processual, não apenas pela reestruturação de institutos, mas, fundamentalmente, pela expressa positivação de uma base principiológica consonante com os preceitos do Estado Democrático de Direito. A exegese da obra do jurista Humberto Theodoro Júnior demonstra que a compreensão do sistema processual moderno exige uma imersão nos princípios que o informam e nas normas fundamentais que o estruturam. Estes dois conceitos, embora interligados, delineiam o que se convencionou chamar de "processo justo", um paradigma que ultrapassa a simples obediência aos ritos para alcançar a concretização da justiça constitucionalmente prometida.

A jornada pela compreensão do sistema processual civil e aqueles princípios que lhe são próprios, em virtude de ser um “ramo de um organismo maior, que é o direito em sua configuração total”, conforme entende Humberto Theodoro Júnior (2025, p. 41), se inicia com os estudos dos princípios que informam e são observáveis em todo o ordenamento jurídico. Aqueles que, tomados como axiomas, seriam universais e alicerçam as leis que regem o processo. São eles o princípio da legalidade, lógico, dialético e político.

O princípio da legalidade, consagrado no artigo 5º, II, da Constituição da República, estabelece que “ninguém será obrigado a fazer ou deixar de fazer alguma coisa, senão em virtude de lei” (Brasil, 1998). Essa salvaguarda assegura a inviolabilidade de direitos fundamentais, ao passo que limita o exercício do poder estatal⁷¹. Na esfera processual, sua aplicação é expressa no artigo 8º do Código de Processo Civil, que impõe ao juiz o dever de "aplicar o ordenamento jurídico" (Brasil, 2015), restringindo a atuação do Poder Judiciário à observância do princípio da legalidade. Destarte, a jurisdição opera sob uma dupla submissão

⁷⁰ Conforme Streck esclarece (2014, p. 429): “(...) qual seria a validade (ou o sentido) de uma hermenêutica jurídica que admitisse ‘qualquer resposta’, enfim, de uma hermenêutica que admitisse, como Kelsen, que a interpretação judicial é um ato de vontade? Qual seria a utilidade uma hermenêutica que admitisse até mesmo múltiplas respostas para um mesmo caso ‘concreto’? Qual seria a razão de ser de uma teoria hermenêutica que admitisse que o direito é aquilo que o ‘intérprete autorizado’ diz que é? Sem medo de errar, nada mais, nada menos, isso seria retornar ao último princípio epocal da metafísica moderna, a vontade do poder (Wille zur Macht). E, em consequência, estar-se-ia a admitir um ‘grau zero na significação’ e, consequentemente, um constante ‘estado de exceção hermenêutico’. A hermenêutica seria, pois, pré-linguística. Mas, já então, não seria mais ‘hermenêutica’! Por isso, a necessidade de existir respostas corretas em Direito”.

à legalidade: o procedimento deve seguir as formas definidas em lei, em obediência ao devido processo legal, e o provimento de mérito deve se fundamentar na lei material correspondente. Contudo, essa submissão não importa uma atividade judicial de repetição da lei, pois a preservação da legalidade exige do julgador o dever de harmonizar a norma infraconstitucional com os ditames constitucionais.

Interligado à legalidade, segundo a doutrina de Theodoro Júnior, emerge o princípio lógico, que demanda uma sustentação racional para todos os atos e decisões das autoridades públicas. No processo, sua máxima expressão é a exigência constitucional de que todas as decisões judiciais sejam fundamentadas, sob pena de nulidade, consoante o artigo 93º, IX, da Constituição da República. A motivação e publicidade das decisões superam o conceito simplista de resposta às alegações das partes, consolidando-se como uma imposição do contraditório e uma exigência político-institucional, pela qual o Judiciário presta contas à coletividade, permitindo o controle social sobre a retidão e os possíveis excessos no exercício do poder pelo magistrado (Júnior, 2025, p. 41-42).

Complementarmente, o princípio dialético admite que a lógica do Direito difere daquela das ciências exatas, tratando-se de uma lógica da razoabilidade, aferida pelo debate e pela argumentação. As leis, por sua natureza programática, apresentam lacunas e imprecisões que o aplicador do direito deve suprir. A tarefa do juiz, portanto, não é afastar-se da lei, mas adequá-la às especificidades do caso concreto, em harmonia com o sistema jurídico vigente. No processo, a dialética se materializa através do contraditório, que assegura uma ampla discussão entre as partes e o juiz, garantindo a todos a participação efetiva na construção da decisão judicial. O juiz, embora comande o processo, não o faz de modo autoritário, pois suas decisões devem ser o resultado dialético do debate, estritamente adstritas ao que foi argumentado e provado nos autos⁷².

A materialização mais significativa da dialética processual civil se revela na moderna concepção do contraditório, positivada nos artigos 7º, 9º e 10º do Código de Processo Civil. Conforme leciona Theodoro Júnior, o princípio evoluiu do dever de audiência bilateral para se consagrar como um verdadeiro direito de participação e influência na construção da decisão judicial. O corolário dessa garantia é o princípio da "não surpresa"⁷³, que veda ao julgador decidir com base em fundamento sobre o qual os litigantes não tiveram a oportunidade prévia de se manifestar, ainda que se trate de matéria apreciável de ofício (Brasil, 2015). Essa

⁷¹ JÚNIOR, Humberto T. Curso de Direito Processual Civil Vol.1 - 66ª Edição 2025. 66. ed. Rio de Janeiro: Forense, 2025. p. 41.

⁷² JÚNIOR, Humberto T. Curso de Direito Processual Civil, p. 42.

⁷³ JÚNIOR, Humberto T. Curso de Direito Processual Civil, p. 96.

garantia se estende aos fatos e aos fundamentos jurídicos, pois, no processo, fato e direito formam um amálgama indissociável na construção do provimento judicial.

Por fim, o princípio político impõe ao magistrado a obrigação de materializar, em suas decisões, as normas, os princípios e os valores que sustentam a estrutura soberana do Estado Democrático de Direito, conforme delineado na Constituição. Assim, a decisão judicial deve, antes de tudo, promover a efetivação dos direitos e princípios fundamentais, aprimorando a interpretação para que o processo se converta em um autêntico instrumento de realização da justiça prometida constitucionalmente (Júnior, 2025, p. 43). Cumpre ao raciocínio, portanto, que a atividade jurisdicional exercida deve perpassar pela análise da incidência e do alcance dos mandamentos constitucionais no caso concreto, os quais servem como parâmetros de interpretação e como normas de aplicação preferencial sobre a legislação infraconstitucional.

Essa vinculação à ordem jurídica, no entanto, impõe ao julgador um roteiro preciso ao se deparar com a incompatibilidade entre a lei e a Constituição. Mesmo quando a norma se revela inconstitucional, não fica o julgador autorizado, só por isso, a criar, *ex novo*, uma regra fora do ordenamento jurídico em vigor. Seu primeiro dever é afastar a lei viciada e procurar no próprio direito positivo uma outra norma que possa atuar em seu lugar⁷⁴. Somente quando a inconstitucionalidade reconhecida deixar um claro vácuo normativo no sistema é que o juiz estará autorizado a conceber uma solução para o caso concreto, valendo-se da analogia, dos costumes e dos princípios gerais de direito. Desse modo, a atividade criativa judicial é subsidiária, operando apenas na hipótese de omissão legal absoluta e não como primeira opção à invalidade da lei.

A evolução desses princípios informativos universais culmina na positivação de um conjunto de garantias que o CPC de 2015 denomina "Normas Fundamentais do Processo Civil". No cerne dessas normas está o princípio do devido processo legal, insculpido no artigo 5º, LIV e LV, da Constituição, e que funciona como um "superprincípio". Ele não se esgota na observância de formalidades que, atribuído de um aspecto substancial, passa a exercer a missão de pacificação social por meio da justa composição do litígio. Exige-se do juiz que não seja apenas a "boca da lei", mas que exerça uma tarefa integrativa, atualizando e adequando a norma aos fatos e valores em disputa. O magistrado complementa a obra do legislador, utilizando critérios éticos e consuetudinários para que o resultado do processo seja verdadeiramente justo, no plano substancial. Assim, o processo se firma como instrumento de efetivação da própria ordem constitucional. Como leciona Humberto Theodoro Júnior (2025):

⁷⁴ JÚNIOR, Humberto T. Curso de Direito Processual Civil, p. 43.

O juiz, enfim, não repete o discurso do legislador; faz nele integrar os direitos fundamentais, não só na interpretação da lei comum, como na sua aplicação ao quadro fático, e, ainda, de maneira direta, faz atuar e prevalecer a supremacia da Constituição. O devido processo legal, portanto, pressupõe não apenas a aplicação adequada do direito positivo, já que lhe toca, antes de tudo, realizar a vontade soberana das regras e dos princípios constitucionais. A regra infraconstitucional somente será aplicada se se mostrar fiel à Constituição. Do contrário, será recusada. E, mesmo quando a lide for resolvida mediante observância da lei comum, o seu sentido haverá de ser definido segundo a Constituição (Júnior, 2025, p. 46-47).

No mesmo sentido, alicerçada no princípio da publicidade do processo, a jurisdição passou a desempenhar uma função de ordem pública, tornando a busca pela verdade um interesse comum entre as partes, o juiz e a sociedade. Embora a verdade real pareça intangível, o compromisso com a busca da "verdade possível" serve como um guia para a atividade jurisdicional, obrigando o juiz a perseguir a veracidade das alegações por meio da atuação *ex officio* na instrução probatória (Júnior, 2025, p. 55-56). O próprio Código de Processo Civil reforça esse compromisso ao estatuir que "ninguém se exime do dever de colaborar com o Poder Judiciário para o descobrimento da verdade" (Brasil, 2015), o que torna evidente o caráter de obra conjunta de todos os seus sujeitos na formação do provimento justo. A decorrência lógica desse postulado é a comunhão da prova: uma vez integrada ao feito, ela se desvincula da parte que a produziu e passa a pertencer ao juízo, devendo ser considerada pelo magistrado em sua análise, em honra ao seu compromisso com a realidade fática.

O princípio da oralidade, por sua vez, é visto como um fator crucial para a concentração dos atos processuais e para a humanização da justiça. Em sua concepção pura, traduz a identidade física do juiz, a concentração da instrução e do julgamento, e a irrecorribilidade das decisões interlocutórias. No direito brasileiro, entretanto, a oralidade é adotada de forma mitigada, pois não se conservou a identidade física do juiz e admite-se o recurso contra um vasto rol de decisões interlocutórias. Não obstante, o Código prestigia a oralidade ao incentivar o contato direto entre os sujeitos processuais, como nas audiências de conciliação e instrução, e na possibilidade de saneamento do processo em audiência colaborativa, reconhecendo que ninguém melhor do que aquele que ouve as partes e testemunhas reúne condições para avaliar a força de seus relatos. Theodoro Júnior ressalta, que para mitigar os efeitos da ausência da identidade física do juiz, o magistrado que assume o processo tem o poder, conferido pelo artigo 370º do CPC, de determinar a repetição das provas orais colhidas por seu antecessor, caso entenda por necessário, assegurando que, na medida do possível, tenha contato pessoal com os elementos probatórios⁷⁵.

⁷⁵ JÚNIOR, Humberto T. Curso de Direito Processual Civil, p. 63-64.

A síntese desses e outros princípios se encontra na noção de um processo legal e justo⁷⁶. Este que, no plano procedimental, deve assegurar o acesso à justiça, o direito de defesa, o contraditório, a paridade de armas, a independência e imparcialidade do juiz, a motivação das decisões e uma duração razoável. No plano substancial, deve proporcionar a tutela efetiva àquele que tem o direito, com base em critérios de equidade e à luz dos princípios constitucionais. Essa busca por justiça não autoriza, contudo, o julgador a sobrepor uma ética pessoal ao direito positivo; a justiça a ser alcançada é aquela que objetivamente corresponde à efetivação das garantias fundamentais previstas na ordem jurídica. O processo justo, nessa concepção, é aquele que outorga aos litigantes a plena tutela jurisdicional dentro dos limites da legalidade, liberdade e igualdade (Júnior, 2025, p. 50).

Essa concepção de processo justo é potencializada por uma interpretação que o próprio Humberto Theodoro Júnior qualifica como “humanizada”, fundada nos princípios da solidariedade e da fraternidade⁷⁷. No entanto, não se pode conceber um processo justo sem que o acesso à justiça não estivesse garantido, e assim o faz o artigo 5º, XXXV, da Constituição, reiterado pelo artigo 3º do CPC, que o estabelece como a porta de entrada para a realização de todos os demais princípios. Esse acesso não se limita ao direito de ter sua pretensão apreciada em juízo, mas compreende o direito a uma tutela jurisdicional efetiva e justa. Para se ver efetivado, contudo, perpassa pelo conjunto de garantias de natureza individuais e estruturais, como o juiz natural, o contraditório, a ampla defesa, a paridade de armas, a assistência aos necessitados, a independência judicial, a motivação das decisões e a duração razoável do processo.

Conforme defende Theodoro Júnior, sob a perspectiva da natureza individual, o acesso à justiça compreende o direito de provocar o Judiciário e obter resposta, contando com um juiz natural e imparcial, a garantia de influir eficazmente na decisão através do contraditório e da ampla defesa, a paridade de armas e a segurança jurídica conferida pela coisa julgada. Já a natureza estrutural do acesso à justiça, exige que o próprio Poder Judiciário se organize com base em garantias como a independência dos juízes, a obrigatoriedade de motivar todas as decisões, a publicidade dos atos, a duração razoável do processo e, acima de tudo, a capacidade de conferir efetividade qualitativa, ou seja, de “dar a quem tem direito tudo aquilo a que faz jus de acordo com o ordenamento jurídico” (Júnior, 2025, p. 72-73).

⁷⁶ Segundo Júnior (2025, p. 49): “Justo e injusto medem-se, no processo, pelos padrões objetivos próprios do direito, e não pela ótica subjetiva e intimista da moral, mesmo porque não é possível na ordem prática quantificar e delimitar, com precisão, os valores e preceitos puramente éticos, em todo seu alcance in concreto”.

⁷⁷ JÚNIOR, Humberto T. Curso de Direito Processual Civil, p. 52-53.

Na aplicação de todo esse arcabouço normativo, o juiz frequentemente se depara com a colisão de princípios. Para solucionar tais conflitos, o artigo 8º do CPC impõe a observância hermenêutica⁷⁸ das máximas da proporcionalidade e da razoabilidade. Tais critérios não são princípios em si, mas postulados normativos aplicativos⁷⁹, ou seja, critérios de raciocínio para resolver antinomias concretas. A razoabilidade atua como um critério de congruência, obstando a aplicação de normas de maneira incompatível com o princípio da igualdade. Já a proporcionalidade aplica-se na análise de uma relação meio-fim, examinando se a medida adotada é (i) adequada para atingir a finalidade, (ii) necessária, ou seja, a menos restritiva possível, e (iii) proporcional em sentido estrito, o que implica uma ponderação entre a importância do fim almejado e a intensidade da restrição a direitos fundamentais.

Por fim, a estrutura do processo justo se completa com os pilares da publicidade e da fundamentação das decisões. A publicidade é uma exigência⁸⁰ do Estado Democrático que permite o controle social sobre a atividade jurisdicional, garantindo a paz e a harmonia social. A ela se integra a obrigatoriedade de fundamentação, sem a qual não é possível aferir se o contraditório foi respeitado nem viabilizar o controle das decisões pelas instâncias superiores. A fundamentação é, ao mesmo tempo, um dever do juiz e um direito da parte de ver seus argumentos considerados. No atual sistema processual, que valoriza os precedentes judiciais, a importância da fundamentação é acentuada, pois é na *ratio decidendi* — o núcleo da fundamentação — que reside a força normativa do precedente. Uma decisão só pode ser considerada racionalmente motivada se sua fundamentação for existente, completa e coerente, enfrentando não apenas os argumentos que a confirmam⁸¹, mas também explicando as razões

⁷⁸ Como observa Humberto Theodoro Júnior (2025, p. 100-101), “a lei nunca se exaure no texto que o legislador lhe deu. Como linguagem, a norma legal, antes de ser aplicada pelo juiz, terá de ser interpretada; e a interpretação, in casu, é ato complexo, pois terá de descobrir o sentido que seja compatível com o sistema normativo total em que a lei se insere; terá, ainda, de considerar o fim visado pelo legislador; e, por último, terá de analisar e encontrar o modo pelo qual a norma abstrata incidirá sobre o quadro fático em que eclodiu o litígio”.

⁷⁹ Esclarece Humberto Ávila (2008, p. 133): “Postulados normativos aplicativos, na medida em que se aplicam para solucionar questões que surgem com a aplicação do Direito, especialmente para solucionar antinomias contingentes, concretas e externas”.

⁸⁰ Segundo Júnior (2025, p. 115-116): “Explica-se a exigência constitucional pela circunstância de que na prestação jurisdicional há um interesse público maior do que o privado defendido pelas partes. Trata-se da garantia da paz e harmonia social, procurada por meio da manutenção da ordem jurídica. Daí que todos, e não apenas os litigantes, têm direito de conhecer e acompanhar tudo o que se passa durante o processo”.

⁸¹ No dizer de Júnior (2025, p. 118-119), “assim como a fundamentação do julgado não pode limitar-se, no plano fático, apenas à avaliação de algumas provas havidas como suficientes para revelação da veracidade da situação definidora da causa, também no plano da identificação da norma aplicável na resolução do conflito (objeto do processo), não pode o julgador trazer apenas argumento de sustentação da norma escolhida como a adequada. É necessário que o ordenamento jurídico seja visto como um todo, que envolve regras legais, princípios e valores, principalmente os consagrados constitucionalmente, e também os que tradicionalmente vigoram no consenso social. (...) O julgamento somente será constitucional e legitimamente fundamentado quando justificar a superioridade do argumento acolhido em relação aos outros eventualmente existentes em sentido diverso, principalmente aquele trazido ao processo pela parte vencida”.

para rechaçar as provas e teses contrárias, especialmente as levantadas pela parte vencida.

4. DA TÉCNICA À CRÍTICA NO HORIZONTE DECISÓRIO

A ascensão das chamadas Inteligências Artificiais Generativas, sobretudo em sua configuração linguístico-discursiva, não há de ser compreendida tão somente como um avanço técnico ou como instrumento de inovação institucional. A emergência desses sistemas inscreve-se, antes, na história da modernidade como um evento de inflexão ontológica, no qual o estatuto do sujeito, do sentido e da normatividade é posto em xeque. Para o presente momento, portanto, não se trata apenas de avaliar se máquinas podem decidir, mas de interrogar o que restaria da decisão quando o sujeito que decide é, ele mesmo, diluído na lógica algorítmica da predição.

Assim, superadas as exposições propedêuticas que, em um primeiro momento, cuidaram de desnudar os fundamentos técnico-operacionais que alicerçam as Inteligências Artificiais Generativas e, em um segundo, exploraram a complexidade filosófico-jurídica do ato de decidir, à luz das matrizes hermenêuticas, axiológicas e normativas que permeiam a atuação jurisdicional, inaugura-se, neste ponto, uma etapa de confronto analítico entre esses dois campos epistêmicos. Pretende-se, com efeito, cartografar a zona de atrito entre a tecnicidade estatística própria das IAGs e a complexa tessitura normativa que sustenta a prática jurisdicional, a fim de inquirir as dissonâncias, identificar os pontos de fratura e aquilatar a compatibilidade — ou a manifesta incompatibilidade — entre a automação da redação decisória e os postulados éticos e jurídicos que alicerçam a atividade jurisdicional em um Estado Democrático de Direito.

De um lado, o primeiro capítulo demonstrou que as IAGs, a despeito de sua impressionante capacidade para gerar discursos fluentes e aparentemente coerentes, operam com base em mecanismos de correlação probabilística, orientados por algoritmos de predição autorregressiva. Tais modelos, como os de arquitetura *Transformer*, são constituídos não por uma inteligência no sentido forte⁸² do termo, mas por uma potência estatística desprovida de intencionalidade ou significado real daquilo que incorpora ou produz. Em verdade, reproduzem padrões linguísticos presentes em vastos *corpora* de dados, mas são ontologicamente inábeis a compreendê-los ou valora-los. O conhecimento que mobilizam é

⁸² Como observa Russel e Norvig (2013, p. 1173), “a asserção de que as máquinas talvez possam agir de maneira inteligente (ou, quem sabe, agir como se fossem inteligentes) é chamada hipótese de IA fraca pelos filósofos, e a asserção de que as máquinas que o fazem estão realmente pensando (em vez de simularem o pensamento) é chamada hipótese de IA forte”.

um retrato estático⁸³ de seu treinamento, incapaz de integrar elementos novos de maneira autônoma, tampouco de dialogar com a singularidade dos eventos que lhe são submetidos. Trata-se, pois, de uma racionalidade técnica, refratária à interpelação e impermeável à historicidade.

Em sentido diametralmente oposto, o segundo capítulo, ao se debruçar sobre a filosofia jurídica de Ronald Dworkin e na perspectiva dos princípios processuais, sistematizados por Humberto Theodoro Júnior, revelou a decisão judicial como expressão densa de responsabilidade institucional, de compromisso com a integridade do Direito e de aderência à dimensão principiológica da jurisdição. A metáfora do juiz Hércules e a premissa do "romance em cadeia" sinalizam que julgar não é um exercício de subsunção, mas um ato de interpretação construtiva, no qual o julgador tem o dever de apresentar a única resposta que confira integridade e coerência à totalidade do Direito. É dizer, a fundamentação de um veredito, nessa concepção, exige uma justificação que se harmonize com o arcabouço jurídico pregresso e, de forma crucial, que seja legitimada pelos princípios de justiça e equidade que o sustentam. Trata-se, em essência, de práxis hermenêutica que compreende, valora e aplica princípios, e não apenas da replicação de padrões textuais sem sensibilidade para as particularidades fático-probatórias irreduzíveis e, por vezes, irrepetíveis do caso concreto.

É precisamente nesse ponto que se situa a dissonância estrutural entre a lógica do processamento computacional e a racionalidade jurídico-hermenêutica. Enquanto os modelos computacionais operam retrospectivamente, recombinao dados segundo regularidades estatísticas, o juiz, na qualidade de operador da integridade, ainda que se valha do passado, o faz com vistas à construção de um futuro coerente, para estender a narrativa do Direito da forma mais justa possível. Ao passo que a máquina reduz o sentido à frequência e prediz, o juiz compreende, aplica e justifica. A primeira calcula correlações; o segundo constrói sentido. Nesse plano, como bem assinala Oliveira (2025), substituir um pelo outro equivale a reduzir o ato de decidir ao “de fazer operar: de ativar uma maquinaria cujo critério não é o justo, o equitativo, o razoável, mas o correlato, o previsível, o performático”.

Acrescente-se, ademais, que a opacidade algorítmica contrasta de modo irreconciliável com a exigência constitucional de fundamentação das decisões judiciais. A impossibilidade de se rastrear com clareza o nexo causal entre os dados de entrada e a saída textual elaborada

⁸³ Conforme completa Arão (2024): “buscando a objetividade e a precisão, acredita-se que as imagens modeladas através de números (como gráficos, tabelas etc.) são a residência da verdade. Tomam-se as equações nas lousas e os algoritmos das inteligências artificiais como a realidade do mundo material. Porém, os números e os símbolos podem oferecer somente representações e idealizações amparadas em abstrações. Isso quer dizer que, na história contada pelos algarismos e algoritmos, não está o mundo real, somente uma representação dele”.

pelo modelo colide frontalmente com a garantia constitucional insculpida no artigo 93, IX, da Constituição da República e 489, §1º, do Código de Processo Civil. A fundamentação, nesse contexto, não se circunscreve apenas aos requisitos de formalidade procedimental, mas representa a condição mesma da legitimidade do exercício do poder jurisdicional e permite o contraditório, a revisão e a responsabilização institucional. Um veredito cujo percurso lógico é inescrutável é, por definição, uma não-decisão no paradigma do Estado Democrático de Direito, pois falha em cumprir com os princípios⁸⁴ que sustentam a juridicidade democrática.

Não se está diante, pois, de um déficit técnico meramente contingente, sanável por aprimoramentos incrementais. Trata-se de um abismo epistemológico entre duas formas incongruentes de produzir sentido: a lógica matemática, reducionista e indiferente à alteridade, e a racionalidade jurídica, que é historicizada, situada e teleológica. Some-se a isso a propensão intrínseca desses sistemas de internalizar e amplificar os vieses do seus dados de treinamento, que, conjugada à sua incapacidade de discernir a veracidade do que produzem, representa uma ameaça direta à segurança jurídica, à isonomia e à busca pela justiça constitucionalmente prometida. À luz desse quadro, a reprodução de vieses compromete, *ad initio*, o que deveria ser a posição equidistante do magistrado em relação às partes, macula a integridade da prestação jurisdicional e atinge, com especial gravidade, os direitos fundamentais daqueles que deveriam ser protegidos por ela.

Nesse cenário, se impõe a transição do encantamento tecnocrático para a crítica: não há de se negar a potencialidade das IAGs⁸⁵ enquanto instrumentos auxiliares, especialmente em tarefas de natureza repetitiva, de pesquisa, sintetização e triagem temática. Todavia, sua adoção em núcleos decisórios do processo, especialmente na redação de sentenças que digam respeito ao mérito e à resolução de controvérsias entre garantias fundamentais, deve ser submetida a um escrutínio rigoroso. A aferição de sua admissibilidade não pode se limitar à eficiência ou à produtividade, mas deve considerar, sobretudo, sua consonância com os princípios estruturantes do Estado Democrático de Direito, sob pena de se instaurar um modelo de jurisdição invertebrado. Isso porque, no domínio jurisdicional, a forma do texto importa menos que o sentido normativo que ele produz e os efeitos jurídicos que daí decorrem.

⁸⁴ Como observam Dinamarco, Badaró e Lopes (2020, p. 109), os princípios são as “ideias mestras externas à ordem processual, geralmente consagradas pela Constituição Federal e responsáveis pela aderência do processo aos grandes valores políticos, sociais, éticos e humanos da sociedade”.

⁸⁵ Conforme Teigão e Fogaça (2025) os “sistemas automatizados permitem acelerar a triagem de demandas, a elaboração de despachos e a movimentação interna de autos, liberando magistrados e servidores para se dedicarem a tarefas de maior densidade jurídica e humana”.

É, pois, a partir dessa tensão constitutiva que as seções seguintes se debruçarão sobre a problemática, investigando, primeiramente, as limitações técnicas que emergem quando a lógica estatística das IAGs colide com a necessidade de uma hermenêutica jurídica e de uma fundamentação transparente. Em seguida, serão explorados os riscos éticos e jurídicos associados à introdução indiscriminada desses sistemas no Judiciário. Por fim, buscar-se-á demonstrar, com base em argumentos de matriz normativa, a desconformidade de uma decisão puramente automatizada com os preceitos fundantes da ordem constitucional brasileira.

4.1. As Fronteiras da Máquina e as Limitações das IAGs na Deliberação Jurisdicional

Diante da complexidade intrínseca à atividade judicante, se impõe ao intérprete um contínuo exercício de reconstrução valorativa, de contextualização hermenêutica e de atenção responsiva às contingências do caso concreto. Tais exigências, inseparáveis da ideia de justiça enquanto prática situada e da jurisdição enquanto exercício ético de poder, excedem, em sua própria estrutura, as capacidades operacionais das Inteligências Artificiais Generativas, cujas limitações técnicas são expressão direta de sua ontologia estatística e de sua epistemologia correlacional.

No plano estrutural, a arquitetura dos *Large Language Models*, que sustenta a performance textual das IAGs, funda-se em mecanismos de reconhecimento de padrões e predição probabilística. As pesquisas no campo da Ciência da Computação demonstram que esse funcionamento, embora sofisticado do ponto de vista computacional, não permite que esses modelos acessem qualquer forma de compreensão substantiva do conteúdo que articulam⁸⁶. A ausência de consciência contextual, de inferência causal e de intencionalidade cognitiva inviabiliza a incorporação dos elementos teleológicos e axiológicos que são inerentes ao Direito e estruturantes da decisão judicial. Nesse sentido, é relevante destacar que o software, enquanto artefato lógico-matemático, não possui qualquer forma de consciência sobre a veracidade, validade ou coerência do conteúdo que gera: seu procedimento inferencial

⁸⁶ A esse respeito, Bender et al. (2021, tradução livre do original), cunharam a expressão “papagaio estocástico” para se referir aos modelos LLM, caracterizando-os como sistemas “para costurar aleatoriamente sequências de formas linguísticas que observou em seus vastos dados de treinamento, de acordo com informações probabilísticas sobre como elas se combinam, mas sem qualquer referência ao significado”. Em linha semelhante, Asher et al. (2023, tradução livre do original), ressaltam as limitações desses modelos quanto aos significados, ao asseverarem que “se você não consegue explorar os significados das palavras em um raciocínio correto, você não sabe realmente o que elas significam”.

é indiferente ao sentido e à correspondência factual da informação⁸⁷. Por conseguinte, inconsistências, obsolescências e imprecisões eventualmente presentes nos dados de treinamento não apenas são absorvidas, mas podem ser inadvertidamente reiteradas nas respostas produzidas.

Em função disso, destaca-se a incapacidade das IAGs de realizar operações de abstração jurídica e de interpretação normativa, especialmente nos casos em que se exige a ponderação entre princípios, a análise das especificidades fáticas e a valoração do impacto social da decisão. Esses elementos, ainda que por vezes possam ser reproduzidos formalmente nos *outputs* gerados pelos modelos, carecem de respaldo hermenêutico autêntico, sendo meras simulações sintáticas de uma racionalidade aberta que lhes é estranha⁸⁸. Não se trata, portanto, de um obstáculo tecnológico superável por avanços incrementais, mas de uma limitação estrutural que decorre da forma como esses sistemas “aprendem” — por estatística e correlação, e não por compreensão ou deliberação.

Não menos relevante é a limitação imposta pela chamada janela de contexto — o intervalo finito de tokens que um modelo é capaz de processar em cada inferência. Esse limite, ainda que tecnologicamente ampliado em versões mais recentes dos modelos, tende a comprometer a possibilidade da IAG de considerar simultaneamente todos os elementos de um processo judicial complexo, como documentos, provas, petições, jurisprudências vinculantes e legislações incidentes. A consequência é a perda de integridade textual, a fragmentação do argumento e a possibilidade concreta de omissões graves ou contradições internas. A decisão, assim construída, se distancia perigosamente da completude e da coerência que o devido processo legal impõe.

A isso se soma o fenômeno já documentado das alucinações algorítmicas — conteúdos espúrios, fictícios ou factualmente incorretos, gerados pela modelagem estatística

⁸⁷ De acordo com os experimentos conduzidos por Dahl et al. (2024), os modelos de linguagem de larga escala tendem a superestimar sistematicamente seu grau de confiança em relação à efetiva taxa de alucinações que produzem. Os autores, nesse contexto, advertem os usuários quanto à necessidade de tomarem as respostas geradas com cautela, uma vez que estas podem ser alucinações em que o modelo está excessivamente confiante e propenso a repetir novamente. Ressaltam, ainda, que embora o GPT-4 se destaque como o modelo menos propenso a alucinações, quando estas ocorrem, tendem a ser particularmente persuasivas, aumentando significativamente o risco de enganar os usuários.

⁸⁸ Conforme destacam Meneceur e Barbaro (2022, tradução livre do original), os modelos encontram óbice na natureza aberta do raciocínio jurídico em virtude de sua dependência ao formalismo matemático, isso porque ele não “permite revelar a complexidade dos argumentos jurídicos utilizados”.

sem qualquer verificação empírica⁸⁹. No âmbito do processo judicial, em que a aderência à verdade material e à legalidade estrita constitui pressuposto da legitimidade do julgado, tal característica representa um risco inadmissível. Uma sentença que incorra em falsificações involuntárias viola frontalmente os princípios da juridicidade, da isonomia e da justiça material. O que se perde, nesse caso, não é apenas a exatidão do conteúdo, mas o próprio sentido do julgar como prática ética e responsiva. É nesse caminho que se demonstram relevantes os exemplos já mencionados das decisões que, em tese, se valeram das GenAI na sua produção, para lançar luz sobre os perigos da confiança excessiva depositada em seus resultados.

Cumpra igualmente destacar a densa opacidade técnica que recobre os mecanismos internos das Inteligências Artificiais Generativas, cuja análise detida será oportunamente empreendida em seção específica. Os LLMs, por se estruturarem com bilhões de parâmetros ajustáveis e com arquiteturas profundas de rede neural, constituem sistemas em que interações internas permanecem, em grande medida, incompreensíveis, inclusive para seus próprios desenvolvedores. A ausência de rastreabilidade compreensível dos critérios decisórios representa uma ameaça concreta à exigência constitucional de fundamentação das decisões judiciais, e reitera a impossibilidade de exercer controle recursal efetivo sobre uma decisão cuja lógica permanece oculta. A ausência de inteligibilidade do percurso argumentativo compromete o contraditório e frustra o direito à ampla defesa.

Por derradeiro, cabe apontar a inércia epistêmica que caracteriza os modelos. Uma vez treinados, os LLMs não se atualizam, tampouco assimilam modificações legislativas, jurisprudenciais ou transformações sociais, salvo mediante novo treinamento. Essa limitação temporal, compreendida por Bender *et al.* (2021) como “value-lock”⁹⁰, impede que acompanhem a historicidade do Direito, que é, por natureza, mutável e sensível aos conflitos que emergem no presente. Um modelo que não acompanha essas mutações tende a cristalizar o passado, reproduzindo seletivamente o que foi, e incapaz de dialogar com o que é — muito

⁸⁹ No esforço de avaliar a aplicabilidade dos LLMs na prática jurídica, Dahl et al. (2024) empreenderam uma análise da recorrência de alucinações produzidas pelos modelos. Os autores constataram uma incidência generalizada desse fenômeno, concluindo que, quando instados a responder perguntas diretas e passíveis de verificação objetiva, formuladas a partir de casos aleatórios da jurisprudência federal norte-americana, os modelos apresentaram respostas alucinatórias em proporções variáveis, oscilando entre 58% (ChatGPT-4) e 88% (Llama 2).

⁹⁰ Segundo os autores, tecnologias baseadas em LLMs tendem a reificar compreensões antigas e excludentes, uma vez que transformações linguísticas e sociais são frequentemente mal assimiladas ou diluídas no volume massivo de dados de treinamento — sobretudo quando não são continuamente atualizados. Nesse sentido, concluem que “considerando apenas os custos computacionais do treinamento de grandes LMs, provavelmente não é viável, mesmo para grandes corporações, retrainá-los completamente com frequência suficiente para acompanhar o tipo de mudança linguística discutida aqui” (Bender et al., 2021, tradução livre do original).

menos com o que pode ser. Assim, a atuação da GenAI reduziria a atividade jurisdicional a uma dimensão estática, fadada à ditadura do passado⁹¹, e alheia à apreciação crítica de contextos sociais que se transformem em paradigmas. Mesmo que, hipoteticamente, fossem capazes de interpretar os paradigmas no sentido proposto por Dworkin, não poderiam enfrentá-los para, assim, criar algo efetivamente novo para o futuro e que caminhasse ao lado do próprio fato social.

Diante dessas considerações, impõe-se uma advertência fundamental: a despeito do fascínio técnico e do entusiasmo institucional que cercam o uso de IAGs, é necessário reconhecer, com sobriedade, que sua inserção na atividade decisória deve ser mantida, no atual estágio, dentro dos limites da função assistiva, voltada a tarefas acessórias e controladas. Qualquer delegação substancial da função jurisdicional a tais sistemas implicaria, inevitavelmente, a corrosão da racionalidade jurídico-democrática que sustenta o Poder Judiciário, convertendo o julgamento em mera operação maquínica desprovida de legitimidade. Preservar a integridade da decisão exige mais do que eficiência — exige humanidade.

4.1.1. A Hermenêutica em Ruínas: A Supressão do Sentido pela Estatística

Não há verdadeira compreensão do fenômeno jurídico sem o reconhecimento de que interpretar é um gesto de responsabilidade ontológica, e não um procedimento técnico de ordenação de dados. A ideia de que sistemas estatísticos possam replicar o ato de julgar encerra, em seu âmago, uma profunda incompreensão da natureza da racionalidade jurídica. Não se evidencia aqui uma simples divergência metodológica entre modelos interpretativos e técnicas computacionais, mas de uma cisão epistemológica profunda entre dois modos de produzir sentido: de um lado, o cálculo, cujo horizonte é a repetição estatística do já dito; de outro, a deliberação, que, orientada por finalidades constitucionais, inaugura sentidos⁹² a partir do inédito.

⁹¹ Segundo Pasquinelli e Joler (2021, p. 1275, tradução livre do original): “(...) como uma técnica de compressão de informações, o aprendizado de máquina automatiza a ditadura do passado, de taxonomias e padrões comportamentais passados, sobre o presente. Esse problema pode ser chamado de regeneração do antigo — a aplicação de uma visão espaço-temporal homogênea que restringe a possibilidade de um novo evento histórico”.

⁹² Chomsky, Roberts e Watumull (2023, tradução livre do original) argumentam que “a mente humana não é, como o ChatGPT e similares, um pesado mecanismo estatístico de correspondência de padrões, que se empanurra com centenas de terabytes de dados e extrapola a resposta mais provável para uma conversa ou a resposta mais provável para uma pergunta científica. Pelo contrário, a mente humana é um sistema surpreendentemente eficiente e até elegante que opera com pequenas quantidades de informação; ela não busca inferir correlações brutas entre pontos de dados, mas sim criar explicações”.

A hermenêutica, enquanto método e postura filosófica diante do texto normativo, compreende a interpretação que considera o contexto histórico-social da norma, os valores constitucionais envolvidos e a finalidade teleológica da aplicação do Direito. Não basta, ao jurista, identificar a literalidade da norma; é preciso perscrutar sua gênese, sua função no ordenamento, seus efeitos no mundo vivido. Nessa chave, a interpretação se funda na ponderação de princípios, na análise da concretude fática e na articulação entre texto, contexto e valor. Sob a teoria da integridade, proposta por Ronald Dworkin, julgar é um gesto construtivo: o juiz deve se comportar como autor de um capítulo coerente em um “romance em cadeia”, capaz de conferir à decisão uma justificabilidade que não apenas se harmonize com a história institucional, mas também traduza o Direito sob sua luz mais justificável, aquela que comungue dos valores de justiça e equidade.

Esse exercício exige a distinção entre regras e princípios, a ponderação destes últimos segundo sua dimensão de peso ou importância e sensibilidade para as singularidades do caso concreto. A subsunção, nesse contexto, não é um silogismo simples, mas o resultado de um contínuo exercício interpretativo que perscruta toda a teia do direito para edificar a solução que melhor expresse a integridade do ordenamento. Não há subsunção que se cumpra em silêncio. O processo decisório, assim compreendido, é indissociável da deliberação e da responsabilidade. A questão final, portanto, parece superar a dimensão restritiva de apresentar respostas às questões e, por óbvio, explicá-las, para alcançar, conforme adverte Streck (2020), a virtude de saber fazer as perguntas.

As IAGs, por outro lado, por mais avançadas que sejam, não interpretam: elas correlacionam. Seu funcionamento repousa sobre a estatística, e não sobre o juízo. Os textos que geram não são respostas interpretativas, mas previsões linguísticas. Sua competência não deriva da compreensão material do Direito, mas da capacidade de identificar regularidades estatísticas e inter-relações semânticas em grandes volumes de dados linguísticos e codificá-las em seus bilhões de parâmetros. Quando uma IAG “aplica” uma norma a um fato, não o faz por discernimento racional ou por sensibilidade normativa, mas por replicação probabilística de padrões previamente internalizados⁹³. O algoritmo não efetua a operação de aplicar uma norma geral e abstrata, devidamente interpretada, a um fato concreto e valorado. O que se enuncia, assim, não é uma decisão, mas uma simulação de decisão — um simulacro sintático de um gesto que, no plano jurídico, deve ser normativo, responsivo e comprometido.

⁹³ Meneceur e Barbaro (2022, tradução livre do original) aprofundam a crítica à lógica da correlação ao advertirem que “os projetistas esperam, na melhor das hipóteses, que a detecção de regularidade nos contextos lexicais consiga produzir os mesmos resultados que o processo real de tomada de decisão que produziu esses dados”.

Essa cisão ontológica entre os paradigmas da hermenêutica e os mecanismos estatístico-correlacionais projeta consequências práticas incontornáveis. Enquanto a hermenêutica busca concretizar a justiça no caso concreto, por meio da mediação entre norma e fato, realizada sob uma racionalidade coparticipativa, dialógico-reflexiva e calcada em causalidades, a correlação estatística limita-se à reprodução formal de padrões discursivos extraídos de regularidades pretéritas. Incapaz de lidar com o inédito, com o dissenso ou com aquilo que escapa à recorrência, o modelo algorítmico tende a reiterar estruturas hegemônicas e silenciar enunciados desviantes. Assim, à promessa de eficiência e sob aparência de neutralidade, a inserção da correlação estatística na esfera judicante tende a contribuir para a cristalização de práticas jurídicas de exclusão⁹⁴, marcadas por seletividades sistêmicas, discriminações e disfunções interpretativas, em dissonância com o horizonte emancipatório e transformador consagrado pela ordem constitucional de 1988.

Ao suprimir a interpretação em favor da predição, corre-se o risco de esvaziar o caráter dialógico do Direito, como condição de possibilidade da juridicidade, comprometendo sua abertura ao dissenso e à autocrítica. A hermenêutica jurídica, por sua natureza, admite a existência de múltiplos sentidos e exige que o intérprete se justifique racionalmente perante os argumentos apresentados pelas partes, em um processo que é, por definição, intersubjetivo. A decisão, assim, não é apenas o fechamento de um conflito, mas a proposição de um sentido que se sustenta diante de uma comunidade. As IAGs, todavia, operam em silêncio: capturam as relações semânticas, mas não as interrogam; imitam a forma da resposta, mas ignoram a densidade da pergunta.

É precisamente nesse ponto que se delinea a fronteira da estatística generativa: sua lógica não suporta o vácuo de sentido, não responde ao inusitado ou ao que está fora da distribuição de treino (Out-of-distribution – OOD)⁹⁵, tampouco assume a responsabilidade da decisão. Ao contrário do intérprete humano, compelido a deliberar mesmo quando a semântica se esgarça e a norma silencia⁹⁶, os sistemas generativos retraem-se, oferecendo uma repetição calculada do já dito. Seu horizonte é o passado textual, não o futuro construtivo. Sua atuação se dá em um plano em que os princípios não pesam, os valores não obrigam e os

⁹⁴ Oliveira (2025) argumenta que a “opacidade do algoritmo, nesse contexto, mostra-se menos como uma falha do sistema do que seu próprio princípio operativo: tornar irrecorrível a decisão, torná-la fato consumado”.

⁹⁵ ZHANG, Lily H.; GOLDSTEIN, Mark; RANGANATH, Rajesh. Understanding Failures in Out-of-Distribution Detection with Deep Generative Models. In: Proceedings of the 38th International Conference on Machine Learning (ICML 2021). PMLR, 2021. p. 12427–12436. Disponível em: <https://proceedings.mlr.press/v139/zhang21a.html>. Acesso em: 7 jul. 2025.

⁹⁶ Consoante ao que determina a Lei de Introdução às normas do Direito Brasileiro em seu artigo 4º (Brasil, 1942), “quando a lei for omissa, o juiz decidirá o caso de acordo com a analogia, os costumes e os princípios gerais de direito”.

efeitos não importam. Ainda assim, para que não se recaia em ilusões do idealismo, é necessário reconhecer que a conclusão e a explicação do julgador humano não é infalível, contudo, como advertem Chomsky, Roberts e Watumull (2023, tradução livre do original), o risco do erro é constitutivo do próprio ato de pensar, pois “para estar certo, deve ser possível estar errado”⁹⁷.

O que se instala com as IAGs é um novo regime de enunciação, em que o critério de validade do discurso já não reside na coerência argumentativa, na autoridade da norma ou na justificação pública, mas na plausibilidade estatística, na eficiência performática e na capacidade de manutenção de fluxos. O modelo que prediz o próximo token passa a ocupar o lugar da consciência que delibera; o cálculo substitui o juízo; a função, o fundamento. Trata-se de um deslocamento estrutural na gramática da decisão, em que a resposta torna-se a reverberação estatística de um passado modelado, esvaziando o núcleo ético-normativo da função jurisdicional.

Por isso, a divergência entre hermenêutica e algoritmo não se reduz à dicotomia entre exegese humana e resposta automática, mas revela a assimetria entre dois paradigmas de racionalidade. O primeiro fundado na razão prática, na responsabilidade institucional e na construção compartilhada do sentido e o segundo, ancorado na mimese estatística de discursos pretéritos, impermeável à crítica e alheio às finalidades substantivas do Direito. A hermenêutica, nesse cenário, deixa de ser apenas uma técnica de leitura e se impõe como condição mesma da juridicidade. Sua substituição por mecanismos probabilísticos representa uma ruptura que compromete a legitimidade da jurisdição, sua dimensão dialética e, em última análise, sua própria humanidade. A esse respeito são as lições de Calamandrei (1960), dignas de menção em destaque:

A verdade é que o juiz não é um mecanismo, não é uma máquina calculadora. É um homem vivo, e sua função de individualizar a lei e aplicá-la ao caso concreto, que, *in vitro*, pode ser representada como um silogismo, é, na realidade, uma operação de síntese que se cumpre misteriosa e calorosamente no cadinho selado do espírito, no qual a mediação e a solda entre a lei abstrata e o fato concreto necessitam, para realizarem-se, da intuição e do ardente sentimento de uma consciência laboriosa. Em suma, o sistema da legalidade, e em geral a teoria normativa do direito, é um esquema didático, útil para os juristas, mas essa decomposição lógica da sentença assemelha-se às análises dos químicos que, embora tenham conseguido individualizar todas as substâncias elementares de que se compõe um organismo vivo, não lograram, contudo, aprisionar em fórmulas a centelha da qual a vida

⁹⁷ Os autores ainda propõem que “a inteligência consiste não apenas em conjecturas criativas, mas também em crítica criativa. O pensamento de estilo humano baseia-se em explicações possíveis e na correção de erros, um processo que limita gradualmente as possibilidades que podem ser consideradas racionalmente” (Chomsky; Roberts; Watumull, 2023, tradução livre do original).

brotou pela misteriosa combinação desses elementos. (Calamandrei, 1960, p. 77-78, tradução livre do original)

Ademais, qualquer projeto de automatização decisória que se pretenda legítimo deve partir da compreensão de que o giro ontológico-linguístico⁹⁸ deslocou a noção de verdade da consciência para a linguagem — e, com isso, tornou insustentável qualquer pretensão de neutralidade técnica descolada da historicidade do Direito. A interpretação jurídica não se presta à predição; ela exige deliberação, contexto e valor. Onde a semântica falhar ou não for o suficiente para resolver uma questão⁹⁹, ou seja, onde houve conflito real de sentidos e exigência de ponderação, a IAG, tão especial em capturar nuances semânticas e tão frágil em interpretar, falhará. Não por erro de cálculo, mas por ausência de mundo.

Assim, a hermenêutica jurídica e a estatística computacional habitam domínios distintos de racionalidade. Confundi-los ou tratá-los como equivalentes é ignorar que, no Direito, compreender não é prever, e que decidir não é apenas escolher¹⁰⁰. É assumir, diante do outro, a responsabilidade pela construção do sentido normativo em face das exigências do caso *sub judice*.

4.1.2. O Véu e o Juízo: A Opacidade Algorítmica Frente ao Dever de Fundamentar

O paradigma de oposição entre a opacidade das arquiteturas algorítmicas que sustentam as Inteligência Artificial Generativa e a exigência constitucional de fundamentação das decisões judiciais, representa, talvez, um dos pontos nevrálgicos da crítica contemporânea à inserção desses sistemas na atividade decisória. De um lado, operam-se sistemas que, a despeito de sua eficácia operacional, ocultam os critérios que orientam a geração de suas respostas em redes profundas de natureza estocástica e não interpretável. De outro, o ordenamento jurídico brasileiro preconiza, impreterivelmente, como condição de validade e

⁹⁸ Streck (2013a, p. 14-17) desenvolve a noção de viragem ontológico linguística, a partir da qual evidencia que o sentido jurídico não se radica na consciência subjetiva do julgador, mas se constitui na linguagem partilhada, situada no horizonte histórico da tradição. O significado não resulta da projeção individual de vontade, mas do fenômeno intersubjetivo que condiciona a própria experiência de mundo, “como algo que produzimos e que é condição de nossa possibilidade de estarmos no mundo”.

⁹⁹ Segundo Maximiliano e Marcaro (2022, p. 27): “Não é possível que algumas séries de normas, embora bem feitas, sintéticas, espelhem todas as faces da realidade (...). Por mais hábeis que sejam os elaboradores de um Código, logo depois de promulgado surgem dificuldades e dúvidas sobre a aplicação de dispositivos bem redigidos. Uma centena de homens cultos e experimentados seria incapaz de abranger em sua visão lúcida a infinita variedade dos conflitos de interesses entre os homens”.

¹⁰⁰ Streck (2013a, p. 107-108) se aprofunda nessa discussão e argumenta que a decisão “não pode ser entendida como um ato em que o juiz, diante de várias possibilidades possíveis para a solução de um caso concreto, escolhe aquela que lhe parece mais adequada. (...) Ora, a decisão se dá, não a partir de uma escolha, mas, sim, a partir do comprometimento com algo que se antecipa. No caso da decisão jurídica, esse algo que se antecipa é a compreensão daquilo que a comunidade política constrói como direito”.

legitimidade da atividade judicante, que nenhuma decisão judicial poderá subsistir sem motivação clara, racional e publicamente justificada, conforme dispõe o art. 93, inciso IX, da Constituição da República e reiterado nos arts. 11 e 489 do Código de Processo Civil.

No centro da controvérsia está o que se convencionou chamar de “efeito caixa-preta”, a opacidade algorítmica que decorre do funcionamento das redes neurais profundas. Nos LLMs, essa opacidade decorre da interação entre bilhões de parâmetros ajustados por processos estocásticos não lineares, o que torna ilegível a lógica interna que vincula os dados de entrada às respostas produzidas. O resultado é que, mesmo quando se conhece a entrada e a saída do sistema, os elos intermediários que compõem o percurso inferencial do sistema não se prestam à reconstrução transparente, inviabilizando a formação de uma cadeia causal clara e compreensível. Na perspectiva de Luís Greco (2020, p. 33-37), a crítica vai além da ausência de qualquer fundamentação, o que o autor denomina “*black box 1.0*”, e alcança o problema ainda mais sutil da “*black box 2.0*”, onde o sistema oferece uma justificativa a *posteriori* sem que se possa garantir que aquelas foram as verdadeiras razões que levaram à conclusão.

Entretanto, a crítica de Greco ao que ele denomina de “sobredimensionado” em relação ao problema da *black box 2.0* parece incorrer em reducionismo. De um lado, sustenta que a distinção entre fundamentação sincera e racionalização insincera é tênue mesmo entre julgadores humanos, e recorre à máxima, e o faz citando Wischmeyer, de que “também o homem é uma *black box* para os outros — e para si mesmo”. De outro, postula que o desenvolvimento de sistemas capazes de documentar cada passo do processo decisório seria suficiente para superar o déficit de justificabilidade, descrevendo, assim, a viabilidade de uma *chain of thought*.

Ocorre, no entanto, que tal analogia entre a opacidade algorítmica e a complexidade da mente humana não resiste a um exame mais apurado. A assimetria entre ambos os paradigmas é de natureza qualitativa. Enquanto o ser humano é dotado de intencionalidade e consciência — atributos que lhe permitem refletir sobre os próprios motivos e reelaborar e justificar os argumentos —, o sistema algorítmico apenas simula tal capacidade por meio de mecanismos estatísticos, sem acessar o sentido normativo ou deliberar em sentido estrito. A pretensa documentação das etapas internas, por sua vez, não escapa ao estatuto de heurística aproximativa, já que, diante da vastidão paramétrica que estrutura os LLMs, qualquer cadeia de raciocínio explicitada não passa de uma representação exógena, e não do encadeamento real dos processos internos.

Ignorar essa distinção ontológica e metodológica entre humanos e arquiteturas estatísticas implica minimizar os riscos implicados na adoção de tecnologias opacas em contextos normativamente exigentes. O fato de que o ser humano também é suscetível a erros ou a racionalizações *ex post* não autoriza a equivalência com dispositivos computacionais cujas operações são, por definição, não transparentes e estruturalmente imunes ao controle hermenêutico. Tratar-se-ia, nesse caso, de uma equivalência imperfeita: ao invés de exigir que o paradigma disruptivo seja posto à prova por seus próprios critérios — transparência e auditabilidade —, opta-se por equipará-lo ao seu contraponto humano, naturalizando deficiências que são sistemáticas e, por vezes, irreversíveis. Mais ainda, a opacidade algorítmica carrega consigo o risco adicional da irreconhecibilidade do erro, já que poderá haver decisões que, por reproduzirem vieses históricos de forma sutil, só revelam sua disfuncionalidade em contextos de reiterada incidência¹⁰¹. Nesses casos, mesmo com a exposição de uma *chain of thought*, não se assegura o acesso às verdadeiras razões do modelo, restando ao operador apenas uma narrativa reconstrutiva que simula — mas não revela — os fundamentos da decisão.

Essa característica da opacidade é particularmente problemática no contexto da jurisdição, pois a decisão judicial não se exaure em sua funcionalidade pragmática — ela é um ato de poder que deve ser publicamente justificado. A fundamentação da decisão, longe de ser uma formalidade procedimental ou ornamento retórico, é o elo entre o exercício da função jurisdicional e as garantias constitucionais. É, portanto, por meio dela que o juiz demonstra que sua decisão não é arbitrária¹⁰², mas está vinculada ao ordenamento jurídico, aos princípios constitucionais e ao conjunto fático-probatório constante dos autos. Além disso, integrada ao princípio da publicidade, a motivação permite o exercício do contraditório, o controle recursal e o escrutínio público da atividade jurisdicional¹⁰³.

¹⁰¹ Larson et al. (2016), em estudo publicado pela ProPublica, analisaram o sistema COMPAS (Correctional Offender Management Profiling for Alternative Sanctions), utilizado por tribunais nos Estados Unidos para estimar o risco de reincidência criminal. Os resultados da pesquisa revelaram evidências substanciais de viés racial nas classificações fornecidas pelo sistema: réus negros que não reincidiram foram quase duas vezes mais propensos a serem classificados como “alto risco” do que réus brancos (45 % vs. 23 %), enquanto brancos que reincidiram foram rotulados como “baixo risco” com maior frequência que negros (48 % vs. 28 %). Mesmo após o controle por variáveis como histórico criminal, idade e gênero, verificou-se que os indivíduos negros possuíam 45% mais chance de receber pontuações de risco mais elevadas — inclusive no risco de reincidência violenta, com chance 77 % maior de classificação elevada.

¹⁰² Como Greco (2020, p. 32) observa, “para uma concepção de direito fundada na ideia de racionalidade, contudo, a fundamentação é que confere juridicidade à decisão, o dado que a diferencia de um mero ato de poder. Nós nos interessamos por fundamentações, nós diferenciamos direito e poder, direito e arbítrio, juiz e oráculo (...)”.

¹⁰³ Segundo Humberto Theodoro Júnior (2025, p. 115): “para a observância do sistema democrático de prestação jurisdicional, não basta divulgar a conclusão do julgado; é indispensável que as razões que a sustentam também sejam explicitadas pelo órgão judicante. Só assim será demonstrado que o contraditório efetivo terá sido realizado e respeitado pelo órgão judicial”.

Ocorre que, em virtude da impossibilidade de reconstituir de modo transparente os critérios internos que levaram a um *output* gerado por uma IAG, se compromete a própria essência da decisão fundamentada. Ainda que o texto produzido por um modelo generativo aparente coerência, clareza e adequação linguística, a incapacidade de rastrear os motivos e o encadeamento lógico que sustentaram sua formulação inviabiliza sua legitimação no plano jurídico. Diferentemente do ser humano, capaz de reconstruir e justificar o percurso argumentativo de sua deliberação, a IAG não possui intencionalidade, nem acesso ao sentido normativo do conteúdo que produz. Sua explicabilidade é, quando muito, uma aproximação retroativa, frequentemente construída por métodos externos ou por camadas adicionais, que se prestam a aproximar um grau mínimo de auditabilidade.

A literatura especializada no campo da *Explainable Artificial Intelligence* tem envidado esforços no sentido de atenuar o problema da opacidade, por meio do desenvolvimento de técnicas voltadas à ampliação da auditabilidade e da interpretabilidade dos modelos. Todavia, conforme já discutido, tais soluções são, em regra, apenas expedientes heurísticos que não oferecem garantias de fidelidade ao encadeamento lógico interno efetivamente percorrido pela arquitetura algorítmica, podendo obscurecer a real complexidade dos mecanismos subjacentes ou, pior, conferir um falso verniz de legitimidade a decisões equivocadas. Soma-se a isso a relação inversamente proporcional entre desempenho e interpretabilidade: quanto mais poderoso e performático é o modelo, mais opaco tende a ser seu funcionamento, o que aprofunda o dilema entre eficiência preditiva e inteligibilidade normativa.

Longe de justificar a abdicação da exigência de explicabilidade, a crescente potência computacional impõe a ampliação da responsabilidade. O aumento da performance não pode vir à custa da transparência. No domínio da jurisdição, essa tensão estrutural deve ser resolvida em favor da inteligibilidade, pois a clareza dos fundamentos não é um requisito procedimental, mas uma exigência de justiça. Modelos generativos podem reproduzir com notável precisão estilística os elementos formais de um pronunciamento judicial, mas sua atuação se restringe a um simulacro retórico, destituído de racionalidade deliberativa. Sob a aparência da decisão, oculta-se uma cadeia estatística hermética, cujo percurso decisório não pode ser logicamente reconstruído ou validamente submetido a escrutínio.

A ausência de inteligibilidade alcança, assim, o núcleo do devido processo legal. A decisão cuja justificação permanece inacessível às partes ou à coletividade não apenas frustra o contraditório e impede o controle, mas dissolve as condições de possibilidade do próprio exercício da jurisdição democrática. Em outras palavras, uma decisão cujo iter lógico

permanece inacessível é, por definição, uma decisão nula¹⁰⁴. A fundamentação exige, mais que coerência formal, a demonstração racional do percurso argumentativo adotado, com a devida apreciação das provas, das normas e dos argumentos jurídicos em disputa. Sem isso, a decisão não pode ser controlada, contestada ou legitimamente acatada.

A atividade de fundamentação judicial deve submeter-se aos critérios da racionalidade. Incumbe ao magistrado a elaboração de um discurso juridicamente coerente, logicamente estruturado e sensível às peculiaridades do caso concreto, de modo a evitar contradições internas, respeitar a ordem de prejudicialidade entre as matérias postas em juízo e revelar, com clareza, os nexos que vinculam as premissas adotadas à conclusão alcançada. É precisamente essa racionalidade que viabiliza o controle público da decisão, permitindo que as partes, os tribunais superiores e a sociedade possam identificar eventuais equívocos, insuficiências argumentativas ou desvios em relação ao arcabouço normativo vigente. A legitimidade da jurisdição, assim, não se sustenta apenas na autoridade formal do julgador, mas na transparência e inteligibilidade do percurso decisório que ele constrói.

Nesse ponto, é fundamental reiterar que a inteligibilidade da decisão não é um atributo procedimental ou desiderato ideal, mas um direito fundamental das partes e da coletividade. A motivação jurisdicional é o instrumento que habilita o controle da legalidade e, ao mesmo tempo, confere sentido democrático à submissão das partes à autoridade judicial. O cidadão só reconhece a legitimidade de um comando quando pode compreender seu fundamento — e, quando necessário, pode exercer o direito de discordar dele¹⁰⁵. A linguagem do Direito é, por essência, uma linguagem dirigida ao outro. Onde há opacidade, instala-se o silêncio institucional; onde se cala o fundamento, rompe-se o elo entre Direito e democracia.

Em última instância, a lógica ininteligível das IAGs mostra-se incompatível com a estrutura normativa do processo jurisdicional. A substituição da motivação racional por cadeias de correlação estatística compromete o ethos da função judicante e colide com os pilares constitucionais da transparência, publicidade e responsabilização. Assim, reafirma-se

¹⁰⁴ Conforme assinala Celso de Mello: “A fundamentação dos atos decisórios qualifica-se como pressuposto constitucional de validade e eficácia das decisões emanadas do Poder Judiciário. A inobservância do dever imposto pelo art. 93, IX, da Carta Política, precisamente por traduzir grave transgressão de natureza constitucional, afeta a legitimidade jurídica da decisão e gera, de maneira irremissível, a consequente nulidade do pronunciamento judicial”. (STF. Habeas Corpus n. 80.892/RJ, Rel. Min. Celso de Mello, Segunda Turma, julgado em 16 out. 2001, DJ 23 nov. 2007).

¹⁰⁵ Conforme as lições de Alexandre de Moraes (2013, p. 1338): “A legitimidade democrática do Poder Judiciário baseia-se na aceitação e respeito de suas decisões pelos demais poderes por ele fiscalizados e, principalmente, pela opinião pública, motivo pelo qual todos os seus pronunciamentos devem ser fundamentados e públicos. A verdadeira, duradoura e incontestável legitimidade do Poder Judiciário será concedida pela opinião pública, pois somente ela é que, em definitivo, consagrará ou rejeitará essa instituição, analisando-a em virtude de sua jurisprudência e de sua atuação perante o Estado (...)”.

que a automação da fundamentação decisória, embora tecnicamente promissora, encontra um limite jurídico intransponível: a exigência de que toda decisão judicial seja inteligível, racionalmente justificável e passível de controle.

4.2. Riscos Ético-Jurídicos da Automação Judicial

A automação de decisões judiciais por meio de Inteligências Artificiais Generativas extrapola os limites das incompatibilidades técnicas e adentra um campo minado de riscos de ordem ética e jurídica, cujas consequências podem abalar os próprios alicerces da administração da justiça. Não se trata de resistência especulativa à um futuro ainda por vir, mas de vícios inerentes à própria natureza desses sistemas, cujos efeitos podem violar garantias fundamentais como a dignidade da pessoa humana¹⁰⁶, a imparcialidade, a isonomia, a segurança jurídica e a própria essência humana do ato de julgar. É neste plano que a tecnicidade cede espaço à normatividade que regem o ato de julgar, e a eficiência cede à primazia da responsabilidade.

No plano ético, o primeiro ponto crítico diz respeito à despersonalização da jurisdição. O Direito, enquanto fenômeno histórico e cultural, está irremediavelmente vinculado à alteridade. A atividade judicante, nesse horizonte, exige o reconhecimento do outro como sujeito de direitos e portador de singularidades irredutíveis, cuja situação concreta reclama consideração contextual, empatia e sensibilidade axiológica¹⁰⁷. O juiz, ao interpretar e aplicar o Direito, deve fazê-lo de modo situado, sensível às circunstâncias concretas, atento à historicidade e permeável à complexidade que se apresenta no caso¹⁰⁸. Ao delegar total ou parcialmente esse papel a um sistema automatizado, corre-se o risco de esvaziar essa dimensão relacional do processo e transformar a jurisdição em mero labor estatístico, insensível às particularidades humanas que estruturam a lide. O resultado é, nesse cenário, a

¹⁰⁶ De acordo com Teigão e Fogaça (2025), o reconhecimento da dignidade da pessoa humana no contexto da Inteligência Artificial impõe “que as tecnologias adotadas jamais reduzam o indivíduo à condição de dado estatístico ou objeto de inferência algorítmica, devendo assegurar tratamento igualitário, não discriminatório e compatível com os direitos fundamentais”.

¹⁰⁷ Nesses termos, segundo Luís Greco (2020, p. 46-47), “a máquina, diversamente do ser humano, não se sabe mortal, efêmera e vulnerável. O juiz-robô sempre será um estranho diante do cidadão como pode o robô decidir sobre a vida de um alguém, se ele literalmente não faz ideia de o que é viver? O juiz-robô significa poder de julgar sem responsabilidade de julgador”.

¹⁰⁸ Como apontam Gaio Júnior e Silva (2023), há determinadas situações que se revelam particularmente temerárias quando submetidas à atuação de sistemas automatizados, como é o caso da valoração da prova, da fixação do quantum indenizatório e da análise do *periculum in mora*. Para os autores, “são situações que ao serem levadas à apreciação do Poder Judiciário demandam dos magistrados certo grau de sensibilidade e bom senso, já que o homem poderia mais facilmente capturar nuances e reconhecer emoções e aspectos morais, questões ainda mais complexas para a IA”.

degradação do ato decisório em um procedimento estéril e apartado de sua finalidade pacificadora.

Uma vez que a há manifesta impossibilidade jurídica de poder sem responsabilidade, exige-se que toda decisão tenha um responsável identificável, capaz de prestar contas pelos fundamentos adotados e pelos efeitos produzidos¹⁰⁹. No entanto, os modelos generativos, desenvolvidos por meio de processos técnico-algorítmicos, calibrados por equipes múltiplas e, por vezes, anônimas, instauram uma zona de indeterminação quanto à autoria decisória. Mesmo quando a assinatura é formalmente preservada, o papel do magistrado pode restringir-se a uma função residual de chancela, diante de decisões produzidas por mecanismos que escapam à sua compreensão integral. Essa ficção técnica, além de inadmissível, dissolve o nexo entre poder e responsabilidade, fragilizando o controle institucional e esvaziando a *accountability* judicial. A figura do magistrado, nesse contexto, deixa de ser o agente central do julgamento para tornar-se um homologador de *outputs* estatísticos.

No plano jurídico, um dos riscos mais incisivos se subsume na reprodução de vieses discriminatórios. Como discutido nos capítulos anteriores, os modelos generativos são treinados com dados que refletem os desequilíbrios de poder, os estigmas culturais e as desigualdades sociais historicamente sedimentadas. Transpostos para o exercício jurisdicional, sob o manto da neutralidade algorítmica¹¹⁰, podem reiterar padrões de exclusão e reforçar estereótipos que deveriam, a rigor, ser desconstruídos pela atuação do Direito. A suposta neutralidade algorítmica, longe de representar uma garantia de imparcialidade, frequentemente oculta as opções valorativas incorporadas nas etapas de modelagem, curadoria e parametrização dos dados. Ainda que se pretenda uma seleção rigorosa dos *datasets* utilizados no treinamento, a atividade curatorial permanece sujeita à filtragem subjetiva, tornando o risco de enviesamento uma constante estrutural e não uma exceção contingente. O resultado é a potencial prolação de decisões que, ao revés de corrigirem iniquidades

¹⁰⁹ Segundo Greco (2020, p. 65-66): “quem exerce poder sobre um outro, tem de assumir a responsabilidade que a isso corresponde, o que supera as capacidades de uma máquina, que não participa da vulnerabilidade e da efemeridade da condição humana. Justiça é algo devemos a cada um de nós reciprocamente, enquanto seres humanos. Não podemos escapar desse dever, escondendo-nos atrás de máquinas. Quaisquer passos nesse sentido devem ser, de uma perspectiva ético-jurídica, (des)qualificados como ‘experimentos proibidos’”.

¹¹⁰ Mesmo sob a perspectiva da neutralidade, Belloso Martín (2015, p. 132, tradução livre do original) adverte que “tanto a IA quanto os sistemas jurídicos especializados pressupõem que é possível confiar em uma visão neutra da representação da realidade (ciências naturais como engenharia, física ou biologia). No entanto, o direito não funciona com leis causais; não há perspectivas neutras ou desinteressadas. O significado da norma só é alcançado por meio de um processo interpretativo. Valores pessoais e dimensões culturais, éticas, sociais e emocionais desempenham um papel decisivo. Isso é uma manifestação da textura aberta da linguagem jurídica”.

históricas, solidificam desigualdades estruturais, representando uma afronta direta ao ideal de isonomia material e justiça substancial.

Como alertam Roque e Santos (2021), em consonância com as reflexões anteriormente delineadas, não se sustenta a crença de que decisões maquínicas seriam mais neutras ou isentas por estarem “livres” da interferência humana. Tal suposição repousa sobre um equívoco epistemológico de fundo: o de que a ausência de deliberação humana direta corresponderia à exclusão da subjetividade. As IAGs, contudo, não constituem instâncias em branco, pelo contrário, são dispositivos impregnados de visões de mundo, vieses implícitos e decisões prévias de engenharia que condicionam, silenciosamente, os contornos de suas respostas. Nessa moldura, a automação não anula a parcialidade — apenas a desloca para camadas invisíveis do código, onde se torna mais difícil de ser rastreada, compreendida e contestada. Suas respostas são sempre construções eivadas, portanto, de parcialidades e ilegalidades ocultas.

À medida que o processo decisório se automatiza em graus cada vez mais elevados, amplia-se o risco de opacificação dos fundamentos que sustentam a decisão judicial, comprometendo a transparência indispensável à sua legitimidade. Isso atinge os princípios do contraditório e da ampla defesa e enfraquece a capacidade das partes de compreender, contestar ou recorrer com eficácia da decisão. A situação se torna ainda mais delicada nos cenários em que o afetado se enquadra nas definições legais de hipossuficiência técnica ou econômica, nos quais este pode sequer ter ciência de que a decisão que o afeta foi proferida com apoio algorítmico (Toledo; Pessoa, 2023), inviabilizando estratégias adequadas de impugnação.

A esse quadro soma-se a preocupação com a cláusula da “não surpresa”, positivada no art. 10 do CPC, cuja função é assegurar que nenhuma decisão seja proferida com base em fundamento sobre o qual a parte não tenha tido a oportunidade de se manifestar. Nesse contexto, a provocação formulada por Streck (2020) adquire relevância singular: “E se o robô, ele mesmo, for surpreendido?”. A indagação ressalta o risco de que a IA, por depender exclusivamente de padrões extraídos do *corpus* de dados que conformou seu treinamento, seja incapaz de reconhecer a pertinência de argumentos ou conjunturas inéditas. Mais grave ainda, o sistema pode relegar ao ostracismo elementos que, embora estatisticamente periféricos, seriam, sob o olhar humano, decisivos à formação do convencimento.

Nesse contexto, há de se resistir à tentação reducionista de outorgar à máquina a incumbência de julgar, já que se trata, por excelência, de um exercício ético de

responsabilidade humana, que envolve ponderações difíceis¹¹¹, a deliberada assunção dos dilemas e, sobretudo, das consequências que deles derivam. A abdicação dessa tarefa em favor de processos automatizados significa, em última análise, a renúncia à complexidade que caracteriza o ideal de justiça, bem como o esvaziamento simbólico da jurisdição enquanto espaço de reconhecimento recíproco¹¹². Nesse horizonte, Streck (2025) propõe a radicalização do debate ao conceber o julgamento humano como um direito fundamental. Não se está diante, aqui, da mera necessidade de garantir um resultado ao final da lide, mas de preservar o próprio processo como espaço de escuta, reconhecimento e justiça construída entre sujeitos. Para Streck, julgar é, em última instância, um ato humano que exige intersubjetividade, sensibilidade histórica e empatia, dimensões que escapam à GenAI, seja qual for a sua capacidade de emular padrões linguísticos.

Acresce-se, a esse cenário, o risco sistêmico de uma jurisprudência moldada por modelos estatísticos, marcada pela padronização dos vereditos, pela supressão da criatividade e pela erosão da margem interpretativa que garante a vitalidade do Direito. A submissão da atividade judicante à lógica algorítmica pode inaugurar um paradigma decisório rigidificado, alicerçado em recorrências passadas e impermeável à inovação hermenêutica exigida pelas transformações sociais em curso. Em vez de fomentar a construção dialógica do Direito, esse modelo de racionalidade instrumental favorece uma cultura de obediência algorítmica, orientada à previsibilidade e à contenção de riscos, em detrimento da justiça material. Essa perspectiva elimina a dimensão humana das interações jurídicas, esvazia o valor do diálogo e a sensibilidade que só o contato humano¹¹³ pode proporcionar (Medina, 2024) e colide com a função democrática do Poder Judiciário, desvirtuando sua vocação primordial: a de assegurar direitos, resistir à homogeneização e preservar a dignidade da experiência singular.

¹¹¹ Maximiliano e Marcaro (2022, p. 72) corroboram essa linha de pensamento ao asseverarem que “os juízes, oriundos do povo, devem ficar ao lado dele, e ter inteligência e coração atentos aos seus interesses e necessidades. A atividade dos pretórios não é meramente intelectual e abstrata; deve ter um cunho prático e humano; revelar a existência de bons sentimentos, tato, conhecimento exato das realidades duras da vida”.

¹¹² Nesse sentido, Teigão e Fogaça (2025) argumentam que a “IA pode servir como ferramenta poderosa de apoio, mas não substitui a escuta, a empatia, a prudência e a ponderação que caracterizam o julgamento verdadeiramente justo”.

¹¹³ Medina (2024) completa afirmando que “Quando entregamos à tecnologia a responsabilidade de analisar e decidir processos jurídicos, transformamos singularidades em padrões estatísticos e tratamos vidas humanas como variáveis manipuláveis por códigos de computador”.

Dessa forma, os riscos éticos e jurídicos da automação do processo decisório não podem ser percebidos tão somente como obstáculos técnicos a serem superados¹¹⁴. Tratam-se, afinal, de questões estruturais que exigem um posicionamento claro do Direito: a proteção da dignidade da pessoa humana, da publicidade dos atos e pronunciamentos jurisdicionais, da motivação das decisões e da responsabilidade institucional devem prevalecer sobre a lógica da eficiência e da celeridade técnica.

4.3. A Dissonância Entre Decisões Automatizadas e a Ordem Constitucional Brasileira

A hipótese de utilização de Inteligências Artificiais Generativas como figura central das decisões judiciais, ainda que impulsionada por argumentos de celeridade e economia processual, alicerçada no princípio da razoável duração do processo, traduz o pensamento de que o Direito é “mais um fato a ser descrito e quantificado pela objetividade científica” (Streck; Jung, 2024). Essa perspectiva, ao tratar a mente humana como um mecanismo decodificável e a linguagem como mera ferramenta de designação, choca-se frontalmente com a tradição hermenêutico-filosófica que fundamenta o modelo constitucional brasileiro. Esta, por sua vez, concebe a jurisdição como uma prática essencialmente linguística, interpretativa e normativa, incompatível com a lógica estatística, opaca e desprovida de semântica que caracteriza as IAGs.

A primeira dimensão da desconformidade manifesta-se no plano da legalidade. Conforme exhaustivamente demonstrado, a Constituição da República, em seu artigo 93, inciso IX, e o Código de Processo Civil de 2015, em seus artigos 11 e 489, §1º, impõe ao magistrado o dever de fundamentar seus pronunciamentos com clareza, enfrentando todos os argumentos deduzidos pelas partes e justificando a opção interpretativa adotada. Essa exigência não é meramente formal, mas material, pois visa assegurar a transparência, a racionalidade e o controle da atividade jurisdicional¹¹⁵. Um sistema que gera sentenças com base em predições

¹¹⁴ Nesse sentido Toledo e Pessoa (2023) concluem que, “ante os riscos concretos das decisões por IA — como a reprodução dos juízos de valor dos desenvolvedores das IA ou a acentuação dos vieses cognitivos envolvidos no processo de tomada de decisão — e a incompatibilidade estrutural entre a forma de processamento da IA e a forma de aplicação do Direito, afigura-se recomendável, ao menos por ora, que não sejam delegadas funções decisórias à IA no âmbito judicial. Mesmo decisões interlocutórias que envolvam juízos de valor — ou seja, que impliquem não apenas o conhecimento da realidade, mas a realização de escolhas segundo os fins eleitos —, submetem-se às limitações apontadas”.

¹¹⁵ A esse respeito Dinamarco, Badaró e Lopes (2020, p. 99) completam: “mais modernamente foi sendo salientada a função política da motivação das decisões judiciais, cujos destinatários não são apenas as partes e o juiz competente para julgar eventual recurso, mas a própria nação como um todo e a opinião popular, com a finalidade de aferir-se em concreto a imparcialidade do juiz e a legalidade e justiça das decisões”.

linguísticas, sem permitir o rastreamento efetivo da cadeia argumentativa, frustra esse dever legal e corrói a integridade do ato decisório, convertendo a sentença em um artefato tecnicamente ilegível e juridicamente opaco.

Do ponto de vista constitucional, a desconformidade se agrava. A decisão judicial automatizada materializa a aspiração de um “realismo jurídico 2.0 ou realismo jurídico *high-tech*”¹¹⁶, que, ao herdar a visão de que o direito é apenas a profecia do que os tribunais farão de fato, se agrega à antiga concepção descritiva ou não normativa do direito. Ao se valer de um instrumento incapaz de interpretar a Constituição e de ponderar princípios, violam-se os postulados do devido processo legal, da dignidade da pessoa humana e da inafastabilidade da jurisdição, que pressupõe um julgamento realizado pelo juiz natural, competente, imparcial e responsável. A IAG, como ferramenta algorítmica, carece de todos esses atributos, operando à margem do *ethos* jurídico que estrutura a jurisdição como função estatal. Roque e Santos (2021) concluem que “seria inconstitucional a tomada de decisões exclusivamente por robôs, sem que suas decisões sejam de alguma forma submetidas à revisão humana, sendo assegurado pela Carta Magna o direito público subjetivo de acesso aos juízes”.

Ainda no plano principiológico, o modelo de julgamento algorítmico conflita com o princípio da igualdade material, ao replicar padrões extraídos de bases de dados enviesadas, marcadas por distorções históricas, exclusões sociais e assimetrias estruturais. Ao reproduzir estatisticamente os estigmas codificados nos dados de treinamento, os modelos generativos tornam-se vetores de reforço da desigualdade sob o disfarce de neutralidade matemática, consequência direta de se reduzir o fenômeno jurídico a dados quantificáveis, ignorando sua dimensão compreensiva e normativa. O que deveria ser isonomia torna-se média estatística; o que se exige como equidade é substituído por recorrência probabilística. Nesse cenário, não se trata apenas de erro técnico, mas de violação direta ao princípio da não discriminação — valor estruturante do sistema de justiça e objeto de compromissos internacionais dos quais o Brasil é signatário.

O princípio da publicidade, por sua vez, também é tensionado pela adoção de sistemas opacos e de difícil auditabilidade. A decisão judicial deve ser passível de controle público, não apenas quanto ao seu conteúdo, mas também quanto ao seu processo de formação. A caixa-preta algorítmica que caracteriza os LLMs inviabiliza esse controle, tornando o processo decisório tecnicamente ilegível para as partes, para os tribunais superiores e para a

¹¹⁶ Para Streck e Jung (2024), a concepção do direito e a tarefa dos jurídicos e teóricos disposta no realismo jurídico 2.0 “não é distinta em relação ao realismo jurídico old school: o direito é aquilo que os tribunais dizem que é, e o papel da doutrina circunscreve-se a analisar as causas que levam a jurisprudência a decidir em determinada direção e, assim, prever os seus resultados”.

sociedade civil. Essa ruptura entre forma e conteúdo subverte o ideal republicano de *accountability* judicial — isto é, a prestação de contas que o magistrado deve à sociedade por suas decisões.

Outro vetor de desconformidade relevante diz respeito à incompatibilidade entre a função jurisdicional e a ausência de sensibilidade normativa por parte dos sistemas de IAG. A decisão judicial, especialmente em contextos de colisão de princípios ou lacunas normativas, exige uma postura deliberativa que mobilize argumentos de justiça, equidade, proporcionalidade e razoabilidade. Tais categorias não são codificáveis nem estatisticamente inferíveis — são produtos de reflexão jurídica situada, que considera a história, a cultura, as desigualdades estruturais e as consequências sociais de determinada decisão. A IAG, por definição, não dispõe dessas capacidades e, portanto, não pode desempenhar com legitimidade a função decisória.

Em um plano mais amplo, a decisão judicial automatizada contraria o próprio modelo de jurisdição adotado pelo Estado brasileiro: um modelo principiológico, comprometido com a máxima efetividade dos direitos fundamentais, com a concretização da justiça material e com a dignidade do jurisdicionado. Ao se valer de dispositivos que operam por semelhança estatística e que carecem de normatividade intrínseca, a jurisdição se afasta de seus fundamentos democráticos e se aproxima perigosamente de uma lógica tecnocrática e desumanizada, a lógica da performance, que reduz o Direito à sua expressão sintática e, matematicamente, semântica e esvazia o seu conteúdo substantivo.

A vista disso, a automação do ato de decidir representa uma regressão a um modelo de juiz-autômato, mero aplicador de silogismos, concepção há muito superada no direito. O sistema processual, alicerçado no princípio do devido processo legal em sua dimensão substancial, exige do magistrado uma tarefa integrativa, na qual ele complementa a obra do legislador, adequando a norma aos fatos e valores em disputa para que o resultado do processo seja verdadeiramente justo. Delegar essa função a um algoritmo que opera por correlação estatística e é incapaz de realizar a valoração axiológica ou a ponderação de princípios é abdicar do papel humanizador que a ordem jurídica confere ao juiz, violando a própria essência da jurisdição.

É certo que a incorporação de sistemas de IAG ao aparato institucional do Judiciário brasileiro constitui uma realidade em marcha, vide o sistema MarIA¹¹⁷, anunciado pelo

¹¹⁷ Medina (2024) adverte que a personificação da marIA não é trivial, pois “dar-lhe o nome de marIA, associá-la à figura materna por excelência, cria um laço simbólico poderoso que desarma críticas e reforça sua aceitação social”. E complementa: “ao contrário do que sugere o nome humanizado marIA, não há cuidado ou empatia em suas operações: há protocolos computacionais que seguem instruções predefinidas”.

Supremo Tribunal Federal em dezembro de 2024, e a conclusão de Toledo e Pessoa¹¹⁸ (2023). Não se trata, portanto, de refutar a presença dessas tecnologias na esfera jurisdicional, mas de delimitar, com acurácia normativa, ética e jusfilosófica, os contornos de sua atuação. Como alertam Toledo e Pessoa (2023), “a condição de os seres humanos figurarem como “supervisores” ou “validadores” das decisões, atos processuais e seus respectivos textos produzidos pelas máquinas pode significar um deslocamento ou até substituição, em que, talvez, a atividade humana se torne secundária na tomada de decisão judicial”. Medina (2024) é contundente ao apontar que “celebrar a automação, sem um olhar crítico para suas consequências, pode sinalizar uma aceitação tácita de um Judiciário que privilegia poucos enquanto desumaniza muitos”.

Com efeito, a automatização de pronunciamentos judiciais, no estado da arte de desenvolvimento das GenAI, sem filtros de controle e sem mediação e supervisão humana, não apenas desafia os limites operacionais das tecnologias atualmente disponíveis, mas afronta diretamente os princípios constitucionais, a legislação processual civil e as balizas que devem nortear o construto de provimentos judiciais. Trata-se, portanto, de uma prática juridicamente insustentável, cuja legitimidade não pode ser presumida com base em argumentos de eficiência, mas deve ser rigorosamente examinada à luz do compromisso ético, principiológico e democrático que estrutura o Estado Democrático de Direito.

¹¹⁸ Em pesquisa conduzida em 2023 com o propósito de aferir o grau de incorporação das inteligências artificiais no âmbito do Poder Judiciário brasileiro, os autores constataram que “(...) ao que tudo indica, a prática de delegar a atividade decisória às máquinas no Judiciário nacional já pode se encontrar em franca realização, sem que as partes e seus advogados ou advogadas sejam cientificadas a respeito, pois os textos das decisões judiciais recebem apenas a assinatura eletrônica do magistrado ou magistrada responsável, desprovida da informação de que houve apoio da máquina para a escolha de solução efetuada e para o respectivo texto apresentado” (Toledo; Pessoa, 2023).

5. CONSIDERAÇÕES FINAIS

A presente investigação teve como escopo delinear, sob enfoque técnico e jurídico, os contornos e, sobretudo, os limites ético-normativos da utilização de Inteligências Artificiais Generativas na automatização do desenvolvimento de decisões judiciais. Partindo de um diagnóstico analítico do funcionamento interno desses sistemas — fundados em modelos estatístico-preditivos de larga escala —, restou evidenciado que, malgrado sua impressionante capacidade de produzir textos aparentemente coerentes, estruturalmente verossímeis e formalmente compatíveis com o léxico jurídico, os algoritmos generativos não são dotados da racionalidade prática, da capacidade hermenêutica e da valoração axiológica que se exigem do ato decisório jurisdicional.

Ao longo da exposição, demonstrou-se que o processo de inferência algorítmica das IAGs repousa, em última análise, sobre estruturas matemáticas que apenas mimetizam a linguagem, mas não a compreendem. Ainda que os LLMs consigam reproduzir com notável fluidez estruturas argumentativas juridicamente aceitáveis, sua atuação permanece adstrita à identificação e recombinação de padrões probabilísticos internalizados no treinamento. Por conseguinte, suas respostas não decorrem de processos interpretativos genuínos, mas da otimização de funções de previsão autorregressiva, orientadas por *datasets* opacos, historicamente condicionados e, não raro, enviesados.

Neste contexto, evidenciou-se que a racionalidade estatística dos sistemas generativos é ontologicamente diversa da racionalidade prática que rege a jurisdição. Enquanto o ato de julgar, conforme delineado no Estado Democrático de Direito, demanda ponderação entre princípios, interpretação constitucionalmente orientada, deliberação crítica e compromisso com os valores democráticos e com a singularidade do caso concreto, as IAGs operam à margem desses requisitos. Trata-se, pois, de uma cisão epistemológica incontornável: a linguagem da lei e da justiça não se reduz ao cálculo, assim como a integridade do Direito não pode ser parametrizada sob os termos da eficiência algorítmica.

Importa destacar, ademais, que os déficits técnicos das IAGs não se manifestam de forma isolada ou compartimentalizada; ao contrário, constituem um complexo interdependente de limitações cuja tentativa de correção pontual pode, paradoxalmente, agravar outros aspectos sensíveis do sistema. A título de exemplo, a busca por filtragem de vieses com novos dados, pode implicar o reforço de padrões da própria seleção ou o esforço por aumentar a explicabilidade algorítmica pode comprometer a performance do modelo. Nesse emaranhado de *trade-offs* técnicos, torna-se evidente que não há soluções neutras nem

ajustes inócuos: qualquer modulação de parâmetros incide sobre a estrutura profunda do sistema, reconfigurando suas respostas e suas implicações normativas. Trata-se, pois, de um campo de tensões em que cada tentativa de otimização carrega consigo o risco de deslocar o problema, e não de resolvê-lo.

A crítica à automatização da decisão judicial, portanto, não repousa sobre uma aversão anacrônica à tecnologia, mas sobre o reconhecimento de que a aplicação do Direito exige mais do que a reprodução formal de discursos plausíveis. Requer responsabilidade epistêmica, sensibilidade institucional, capacidade crítica e fundamentação racional — elementos todos que escapam à lógica computacional, mesmo nos modelos mais avançados de GenAI. A opacidade algorítmica, a ausência de *accountability*, os riscos de reprodução de vieses estruturais e a impossibilidade de fundamentação genuinamente inteligível constituem barreiras intransponíveis à sua adoção indiscriminada no núcleo da atividade jurisdicional.

Diante disso, conclui-se que a utilização de IAs generativas pode encontrar espaço legítimo no sistema de justiça apenas em funções auxiliares, voltadas à otimização procedimental ou à organização informacional. No entanto, a substituição do agente humano na formação do juízo — núcleo do ato de julgar — não apenas transgredir os princípios constitucionais da jurisdição, como compromete a legitimidade democrática das decisões judiciais. Julgar é mais do que escolher: é assumir responsabilidade pública sobre a norma aplicada, escutar as partes, interpretar com integridade e produzir sentido normativo situado. A automatização desse processo, se desprovida de uma regulação crítica, pode converter a jurisdição em um simulacro de justiça, esvaziando o papel do juiz como garantidor de direitos e guardião da Constituição. Em última instância, a autoridade do Direito reside na humanidade de sua aplicação — e essa, por definição, não é outorgável a um algoritmo.

Desse modo, um modelo de governança tecnológica compatível com os marcos do Estado Democrático de Direito deve assentar-se sobre alguns princípios estruturantes: (i) transparência, mediante documentação auditável sobre o funcionamento promenorizado dos sistemas utilizados, inclusive com a previsão de mecanismos de *explainability* que possibilitem a reconstrução dos critérios decisórios; (ii) responsabilidade institucional, com a definição de cadeias claras de atribuição e supervisão humana obrigatória em qualquer etapa que envolva juízo, interpretação e decisão; (iii) respeito à integridade epistêmica da função jurisdicional, vedando-se a substituição do julgador por sistemas generativos e restringindo seu uso a funções auxiliares supervisionadas; e (iv) efetividade dos direitos fundamentais, com ênfase na prevenção da discriminação algorítmica e na tutela do contraditório, da ampla defesa e da motivação racional das decisões.

REFERÊNCIAS

ARÃO, Cristian. Por trás da inteligência artificial: uma análise das bases epistemológicas do aprendizado de máquina. Trans/form/ação: Revista de Filosofia da Unesp, Marília, v. 47, n. 2, e02400163, 2024. Disponível em: <https://revistas.marilia.unesp.br/index.php/transformacao/article/view/15196>. Acesso em: 23 fev. 2025.

ASHER, Nicholas *et al.* *Limits for learning with language models*. In: *JOINT CONFERENCE ON LEXICAL AND COMPUTATIONAL SEMANTICS (STARSEM)*, 12., 2023, Toronto. Anais [...]. Toronto: Association for Computational Linguistics, 2023. p. 236–248. DOI: 10.18653/v1/2023.starsem-1.22. Disponível em: <https://aclanthology.org/2023.starsem-1.22/>. Acesso em: 6 jul. 2025.

ATREY, Pranjal; BRUNDAGE, Michael P.; WU, Min; DUTTA, Sanghamitra. *Demystifying the Accuracy-Interpretability Trade-Off: A Case Study of Inferring Ratings from Reviews*. arXiv:2503.07914, 2025. Disponível em: <https://arxiv.org/abs/2503.07914>. Acesso em: 29 jun. 2025.

ÁVILA, Humberto. Teoria dos princípios. 8. ed. São Paulo: Malheiros, 2008.

BARBOSA, Jardeson Leandro Nascimento *et al.* Introdução ao Processamento de Linguagem Natural usando Python. In: Livro Anais - Artigos e Minicursos, v. 1, n. 1, p. 336-360, jun. 2017. III Escola Regional de Informática do Piauí, 2017. Disponível em: https://www.facom.ufu.br/~wendelmelo/terceiros/tutorial_nltk.pdf. Acesso em: 15 jul. 2024.

BARREDO ARRIETA, Alejandro. *et al.* *Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI*. *Information Fusion*. v. 58, p. 82–115, 2020. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S1566253519308103>. Acesso em: 27 mai. 2025.

BENDER, Emily M. *et al.* *On the dangers of stochastic parrots: Can language models be too big?*. *Proceedings of the ACM Conference on Fairness, Accountability, and Transparency*, p. 610-623, 2021. DOI: <https://doi.org/10.1145/3442188.3445922>.

BENDER, Emily M.; KOLLER, Alexander. *Climbing towards NLU: On meaning, form, and understanding in the age of data*. In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 58., 2020, Online. Online: Association for Computational Linguistics, 2020. p. 5185–5198. DOI: 10.18653/v1/2020.acl-main.463. Disponível em: <https://aclanthology.org/2020.acl-main.463/>. Acesso em: 20 jun. 2025.

BELLOSO MARTÍN, Nuria. *Algunas reflexiones sobre la informática jurídica decisional*. In: BAEZ, Narciso Leandro Xavier *et al.* O impacto das novas tecnologias nos direitos fundamentais. Joaçaba: Ed. Unoesc, 2015.

NUNES, Dierle José Coelho; PEDRON, Flávio Barbosa Quinaud; BAHIA, Alexandre Gustavo Melo Franco de Moraes. Precedentes e a busca de uma decisão correta. In: THEODORO JÚNIOR, Humberto (coord.). Processo civil brasileiro: novos rumos a partir do CPC/2015: atualizado de acordo com a Lei 13.256/2016. Organização de Fernando Gonzaga Jayme, Gláucio Maciel Gonçalves, Juliana Cordeiro de Faria, Marcelo Veiga Franco, Mayara de Carvalho Araújo e Suzana Santi Cremasco. Belo Horizonte: Del Rey, 2016. p. 189-199.

BRAGANÇA, Fernanda; BRAGANÇA, Laurinda Fátima da F. P. G. Revolução 4.0 no Poder Judiciário: levantamento do uso de inteligência artificial nos Tribunais brasileiros. Revista da Seção Judiciária do Rio de Janeiro, Rio de Janeiro, v. 23, n. 46, jul./out. 2019. p. 69. DOI: <https://doi.org/10.30749/2177-8337.v23n46p65-76>.

BRASIL. Constituição da República Federativa do Brasil de 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 20 fev. 2025.

BRASIL. Lei nº 13.105, de 16 de março de 2015. Código de Processo Civil. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/L13105.htm. Acesso em: 25 fev. 2025.

BRASIL, Decreto – Lei nº. 4.657, de 4 de setembro de 1942 – Lei de Introdução às Normas do Direito Brasileiro. Disponível em http://www.planalto.gov.br/ccivil_03/decretolei/Del4657compilado.htm. Acesso em: 25 fev. 2025.

BROWN, Tom *et al.* *Language models are few-shot learners. Advances in Neural Information Processing Systems*, v. 33, p. 1877–1901, 2020.

CALAMANDREI, Piero. *Proceso y Democracia*. Tradução de Héctor Fix-Zamudio. Buenos Aires, Ed. Juridicas Europa-america, 1960.

CARVALHO NETTO. Menelick de. A Hermenêutica Constitucional sob o paradigma do Estado Democrático de Direito. In: CATTONI DE OLIVEIRA, Marcelo Andrade (Org.). *Jurisdição e Hermenêutica Constitucional no Estado Democrático de Direito*. Belo Horizonte: Mandamentos, 2004.

CHOMSKY, Noam; ROBERTS, Ian; WATUMULL, Jeffrey. *The false promise of ChatGPT*. *New York Times*, 8 mar. 2023. Disponível em: <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html>. Acesso em: 05 jul. 2025.

CORMEN, Thomas H. *et al.* *Introduction to algorithms*. 3. ed. Cambridge: The MIT Press, 2009. Disponível em: https://enos.itcollege.ee/~japoia/algorithms/GT/Introduction_to_algorithms-3rd%20Edition.pdf. Acesso em: 20 mai. 2025.

COELHO, Inocêncio Mártires. *Interpretação constitucional*. 4ª ed. São Paulo: Saraiva, 2011.

CONSELHO NACIONAL DE JUSTIÇA. Relatório de pesquisa: O uso da Inteligência Artificial Generativa no Poder Judiciário Brasileiro. 2024. Conselho Nacional de Justiça. Brasília: CNJ, 2024. Disponível em:

<https://www.cnj.jus.br/wp-content/uploads/2024/09/cnj-relatorio-de-pesquisa-iag-pj.pdf>. Acesso em: 01 jan. 2025.

DAHL, Matthew *et al.* *Large legal fictions: profiling legal hallucinations in large language models.* *Journal of Legal Analysis*, v. 16, n. 1, p. 64-93, 2024. DOI: <https://doi.org/10.1093/jla/laae003>.

DAMÁSIO, António R. O erro de Descartes: emoção, razão e cérebro humano. Trad. Dora Vicente e Georgina Segurado. São Paulo: Companhia das Letras, 2012.

DINAMARCO, Cândido Rangel; BADARÓ, Gustavo Henrique; LOPES, Bruno Vasconcelos Carrilho. Teoria Geral do Processo. 32. ed. rev. e atual. São Paulo: Malheiros, 2020.

DIREITO, Carlos Alberto Menezes. A decisão judicial. Revista do Tribunal Regional do Trabalho da 15ª Região, Campinas, n. 16, p. 131-146, 2001.

DOSHI-VELEZ, Finale; KIM, Been. *Towards A Rigorous Science of Interpretable Machine Learning.* arXiv:1702.08608, v. 2, n. 2, 2017. Disponível em: <https://arxiv.org/abs/1702.08608>. Acesso em: 27 mai. 2025.

DWORKIN, Ronald. Levando os direitos a sério. Trad. Nelson Boeira. São Paulo: Martins Fontes, 2002.

DWORKIN, Ronald. O império do direito. Trad. Jefferson Luiz Camargo. 3. ed. São Paulo: Martins Fontes, 2014.

FRANCESCHELLI, Giorgio; MUSOLESI, Mirco. *On the creativity of large language models.* *AI & Society*, p. 1-11, 2024. Springer. Disponível em: <https://link.springer.com/article/10.1007/s00146-024-02127-3>. Acesso em: 27 mai. 2025.

FRIEDENBERG, Jay D.; SILVERMAN, Gordon W. *Cognitive Science: An Introduction to the Study of Mind.* 3. ed. Thousand Oaks: SAGE Publications, 2006.

GAIO JÚNIOR, Antônio. P.; SILVA, Fábiana A. Direito, processo e Inteligência Artificial. Diálogos necessários ao exercício da jurisdição. Revista Eletrônica de Direito Processual - REDP, 24(1), 60-99, 2023. Disponível em: <https://www.e-publicacoes.uerj.br/index.php/redp/article/view/72240/44601>. Acesso em 15 jul. 2025.

GARCIA, Ana Cristina Bicharra. Ética e Inteligência Artificial. Computação Brasil, n. 43, p. 14-22, 2020. Disponível em: <https://sol.sbc.org.br/journals/index.php/comp-br/article/view/1791>. Acesso em: 12 mai. 2025.

GOOGLE RESEARCH. Beyond language, vision and generative models. Google Research Blog, 19 dez. 2022. Disponível em: <https://research.google/blog/google-research-2022-beyond-language-vision-and-generative-models/>. Acesso em: 29 mai. 2025.

GRECO, Luís. Poder de julgar sem responsabilidade de julgador: a impossibilidade jurídica do juiz-robô. São Paulo: Marcial Pons, 2020.

IBM. O que é IA generativa?. IBM *Think*, 2024. Disponível em: <https://www.ibm.com/br-pt/think/topics/generative-ai>. Acesso em: 14 mai. 2025.

IBM. *What is a neural network?*. IBM *Think*, 2021. Disponível em: <https://www.ibm.com/think/topics/neural-networks?>. Acesso em: 10 mai, 2025.

JÚNIOR, Humberto T. Curso de Direito Processual Civil Vol.1 - 66ª Edição 2025. 66. ed. Rio de Janeiro: Forense, 2025.

KOUBAA, Anis. GPT-4 vs. GPT-3.5: *A concise showdown*. *TechRxiv*. 2023. DOI: <https://doi.org/10.36227/techrxiv.22312330.v1>.

LARSON, Jeff *et al.* *How We Analyzed the COMPAS Recidivism Algorithm*. ProPublica, 23 maio 2016. Disponível em: <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>. Acesso em: 17 jul. 2025.

MARANHÃO, Juliano. A importância da inteligência artificial inteligível no Direito. JOTA, São Paulo, 22 fev. 2019. Disponível em: <https://www.jota.info/artigos/a-importancia-da-inteligencia-artificial-inteligivel-no-direito-22-022019>. Acesso em: 16 fev. 2025.

MAXIMILIANO, Carlos; MARCARO, Alysso. Coleção Fora de Série - Hermenêutica e Aplicação do Direito - 23ª Edição 2022. 23. ed. Rio de Janeiro: Forense, 2021.

MEDINA, Damares. marIA: tecnologia, opacidade e o futuro da jurisdição constitucional: os desafios de uma revolução algorítmica no Judiciário. JOTA, São Paulo, 23 dez. 2024. Disponível em: <https://www.jota.info/artigos/maria-tecnologia-opacidade-e-o-futuro-da-jurisdicao-constitucional>. Acesso em: 05 jul. 2025

MENECEUR, Yannick; BARBARO, Clementina. *Artificial intelligence and the judicial memory: the great misunderstanding*. *AI Ethics*, v. 2, p. 269–275, 2022. DOI: <https://doi.org/10.1007/s43681-021-00101-z>.

MORAES, Alexandre. Constituição do Brasil interpretada e legislação constitucional. 9ª ed. São Paulo: Atlas, 2013

OLIVEIRA, Priscila Ricardo de. Subjetividade, autoria e processo decisório: Dilemas do Direito frente aos avanços da Inteligência Artificial Generativa. Revista Jurídica Gralha Azul - TJPR, [S. l.], v. 1, n. 28, 2025. Disponível em: <https://revista.tjpr.jus.br/gralhaazul/article/view/195>. Acesso em: 28 mai. 2025.

OPENAI. GPT-4 *Technical Report*. 2023. Disponível em: <https://arxiv.org/html/2303.08774v6>. Acesso em: 15 mai. 2025.

OPENAI. *Extracting Concepts from GPT-4*. 2024. Disponível em: <https://openai.com/index/extracting-concepts-from-gpt-4/>. Acesso em: 15 mai. 2025.

PASQUINELLI, Matteo; JOLER, Vladan. *The Nooscope manifested: AI as an instrument of knowledge extractivism*. *AI & Soc*, v. 36, p. 1263-1280, 2021. Disponível em: <https://doi.org/10.1007/s00146-020-01097-6>. Acesso em: 05 jul. 2025

ROQUE, André Vasconcelos; SANTOS, Lucas Braz Rodrigues dos. Inteligência artificial na tomada de decisões judiciais: três premissas básicas. *Revista Eletrônica de Direito Processual – REDPRio de Janeiro*. Ano 15. Volume 22. Número 1. Janeiro a Abril de 2021.

RUDIN, Cynthia. *Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead*. *Nature machine intelligence*, v. 1, n. 5, p. 206-215, 2019.

RUSSEL, Stuart; NORVIG, Peter. *Inteligência artificial*. Trad. Regina Célia Simille de Macedo. 3. ed. Rio de Janeiro: Elsevier, 2013.

SALVAGNO, Michele; TACCONE, Fabio Silvio; GERLI, Alberto Giovanni. *Can artificial intelligence help for scientific writing? Critical Care*, v. 27, n. 75, 25 fev. 2023. Disponível em: <https://ccforum.biomedcentral.com/articles/10.1186/s13054-023-04380-2>. Acesso em: 12 mai. 2025.

SEARLE, John R. *Mente, cérebro e ciência*. 4ª ed. Tradução de Arthur Morão. Lisboa: Edições 70, 1997.

SENO, Eloize Rossi Marques; CLARO, Daniela; MOTA, Laila; RODRIGUES, Jessica. *Semântica distribucional*. In: CASELI, H. M.; NUNES, M. G. V. (Org.). *Processamento de linguagem natural: conceitos, técnicas e aplicações em português*. 2. ed. BPLN, 2024. Cap. 10. Disponível em: <https://brasileiraspln.com/livro-pln/2a-edicao/parte-significado/cap-semantica-distribucional/cap-semantica-distribucional.html>. Acesso em: 22 mai. 2025.

SHARDLOW, Matthew; PRZYBYŁA, Piotr. *Deanthropomorphising NLP: Can a language model be conscious?* *PLoS ONE*, v. 19, n. 12, p. e0307521, 2024. Disponível em: <https://doi.org/10.1371/journal.pone.0307521>. Acesso em: 1 jul. 2025.

STRECK, Lenio Luiz; JUNG, Luã. *Hermenêutica e Inteligência Artificial: Por uma Alternativa Paradigmática ao Imaginário Técnico-Jurídico*. *Direito Público*, 21(110), 2024. <https://doi.org/10.11117/rdp.v21i110.7689>.

STRECK, Lenio Luiz. *O que é isto - decido conforme minha consciência?* - 4. ed. rev. Porto Alegre: Livraria do Advogado Editora, 2013a.

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. *O(a) robô Maria, precedentes e o direito a um julgamento humano*. Consultor Jurídico, São Paulo, 13 fev. 2025. Disponível em: <https://www.conjur.com.br/2025-fev-13/oa-robo-maria-precedentes-e-o-direito-a-um-julgamento-humano/>. Acesso em: 05 jul. 2025

STRECK, Lenio Luiz. Um robô pode julgar? Quem programa o robô? Consultor Jurídico, São Paulo, 03 setembro de 2020. Disponível em: <https://www.conjur.com.br/2020-set-03/senso-incomum-roboto-julgar-quem-programa-roboto/>. Acesso em: 2 mai. 2025.

STRECK, Lenio Luiz. Why judicial discretion is a problem to Dworkin but not to Alexy/Porque a discricionariedade é um grave problema para Dworkin e não o é para Alexy. *Direito e Práxis*, v. 4, n. 2, p. 343-368, 2013b.

TEIGÃO, Rafael Coninck; FOGAÇA, Leonardo de Andrade Ferraz. O uso ético e responsável da Inteligência Artificial no judiciário brasileiro: impactos para o jurisdicionado e balizas regulatórias. *Revista Jurídica Gralha Azul - TJPR*, [S. l.], v. 1, n. 28, 2025. DOI: 10.62248/w1y8bv17. Disponível em: <https://revista.tjpr.jus.br/gralhaazul/article/view/189>. Acesso em: 13 jul. 2025.

TRINDADE, Alessandra Stefane; OLIVEIRA, Henry Poncio. Inteligência artificial (IA) generativa e competência em informação: habilidades informacionais necessárias ao uso de ferramentas de IA generativa em demandas informacionais de natureza acadêmico-científica. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 29, p. e-47485, 2024. Disponível em: <https://www.scielo.br/j/pci/a/GVCW7KbcRjGVhLSrmy3PCng/?lang=pt#>. Acesso em: 25 fev. 2025.

TOLEDO, Cláudia; PESSOA, Daniel. O uso de inteligência artificial na tomada de decisão judicial. *Revista de Investigações Constitucionais*, v. 10, n. 1, 2023. Disponível em: <https://www.scielo.br/j/rinc/a/qRC4TmVXVDJ8Wkv7Ns49jxH/>. Acesso em: 20 fev. 2025.

VASWANI, Ashish *et al.* *Attention is all you need*. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*. 6000–6010, 2017. Disponível em: https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf. Acesso em: 15 ago. 2025.

WALTER, Yoshija; ZBINDEN, Lukas. *The problem with AI consciousness: a neurogenetic case against synthetic sentience*. arXiv:2301.05397, 2022. Disponível em: <https://arxiv.org/abs/2301.05397>. Acesso em: 1 jul. 2025.

WOLFE, Cameron R. *Decoder-Only Transformers: The Workhorse of Generative LLMs*. *Deep (Learning) Focus*, 2024. Disponível em: <https://cameronrwolfe.substack.com/p/decoder-only-transformers-the-workhorse>. Acesso em: 21 mai. 2025.