

CENTRO UNIVERSITÁRIO AUTÔNOMO DO BRASIL
PROGRAMA DE PÓS-GRADUAÇÃO EM DIREITOS FUNDAMENTAIS
E DEMOCRACIA – DOUTORADO

SÉRGIO RODRIGO DE PÁDUA

DA JURISDIÇÃO “EX MACHINA” AO JUIZ CIBORGUE: O AUXÍLIO
DA INTELIGÊNCIA ARTIFICIAL NA INTERPRETAÇÃO DO DIREITO
NO BRASIL

CURITIBA

2023

CENTRO UNIVERSITÁRIO AUTÔNOMO DO BRASIL
PROGRAMA DE PÓS-GRADUAÇÃO EM DIREITOS FUNDAMENTAIS
E DEMOCRACIA – DOUTORADO

SÉRGIO RODRIGO DE PÁDUA

DA JURISDIÇÃO “EX MACHINA” AO JUIZ CIBORGUE: O AUXÍLIO
DA INTELIGÊNCIA ARTIFICIAL NA INTERPRETAÇÃO DO DIREITO
NO BRASIL

Tese apresentada como requisito à obtenção de título de Doutor, Programa de Pós-Graduação em Direitos Fundamentais e Democracia - Doutorado, Centro Universitário Autônomo do Brasil.

Orientador: Prof. Dr. Bruno Meneses Lorenzetto

CURITIBA

2023

ATA DE DEFESA PÚBLICA DE TESE DE DOUTORADO

Aos 20 dias do mês de março de 2023, às 14h, na Sala 04, PPGD UniBrasil, foi realizada Defesa Pública de Tese do Doutorando **SÉRGIO RODRIGO DE PÁDUA**, no Programa de Pós-Graduação em Direito do Centro Universitário Autônomo do Brasil - UniBrasil, Área de Concentração Direitos Fundamentais e Democracia, Linha de Pesquisa: Jurisdição e Democracia, com o trabalho intitulado: "**DA JURISDIÇÃO "EX MACHINA" AO JUIZ CIBORGUE: O AUXÍLIO DA INTELIGÊNCIA ARTIFICIAL NA INTERPRETAÇÃO DO DIREITO NO BRASIL**" orientado pelo PROFESSOR DR. BRUNO MENESES LORENZETTO. A Banca Examinadora foi constituída pelos PROFESSORES DOUTORES: BRUNO MENESES LORENZETTO (PRESIDENTE), FABIANO HARTMANN PEIXOTO, JOSÉ RENATO GAZIERO CELLA, PAULO RICARDO SCHIER E OCTAVIO CAMPOS FISCHER (MEMBROS).

Resultado final atribuído ao candidato (Art. 105 – Regimento Interno do Programa de Pós-Graduação em Direito do UniBrasil):

- () Reprovado
- () Aprovado com exigências
- () Aprovado
- (☒) Aprovado com láurea

Aprovação com exigências (Art. 105 – Regimento Interno do Programa de Pós-Graduação em Direito do UniBrasil):

Eu, PROF. DR. BRUNO MENESES LORENZETTO, Presidente da Banca e Orientador do Projeto, lavrei a presente Ata que segue assinada por mim e pelos demais membros da Banca Examinadora.

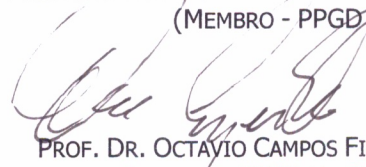


PROF. DR. BRUNO MENESES LORENZETTO
(PRESIDENTE - PPGD UNIBRASIL)



PROF. DR. JOSÉ RENATO GAZIERO CELLA
(MEMBRO - PPGD IMED)

PROF. DR. FABIANO HARTMANN PEIXOTO
(MEMBRO - PPGD UNB)



PROF. DR. OCTAVIO CAMPOS FISCHER
(MEMBRO - PPGD UNIBRASIL)



PROF. DR. PAULO RICARDO SCHIER
(MEMBRO - PPGD UNIBRASIL)

Pádua, Sérgio Rodrigo de
Da Jurisdição "Ex Machina" ao Juiz Ciborgue: o auxílio
da Inteligência Artificial na interpretação do Direito no
Brasil. / Sérgio Rodrigo de Pádua. - Curitiba, 2023.
395 f.

Orientador: Prof. Dr. Bruno Meneses Lorenzetto
Tese (Doutorado) - UniBrasil, 2023.

1. Inteligência Artificial. 2. Atividade
Jurisdicional. 3. Teoria do Direito. 4. Argumentação
Jurídica. I. Lorenzetto, Bruno Meneses, orient. II.
Título.

Dedico esta obra ao meu filho Lucas, brilhante
luz da minha vida.

AGRADECIMENTOS

A pesquisa científica é instigada por um espírito semelhante ao de *Jornada nas Estrelas*, indo aonde nenhuma pessoa jamais esteve. Esse caminho é feito com alegrias e, ao mesmo tempo, com muito suor e intensos estudos, e nunca é percorrido sozinho. Pesquisar é um verbo que deve ser conjugado no plural.

Dessa forma, agradeço primeiramente ao meu orientador, Professor Bruno Meneses Lorenzetto, pela assertividade técnica aliada à elegância ao tratar dos temas mais complexos, pelas correções criteriosas que fizeram a tese se desenvolver, bem como pelo profissionalismo que nos trouxe até a amizade.

Registro os valiosos apontamentos do Professor Paulo Ricardo Schier, que, com a paciência dos grandes mestres, contribuiu para a melhoria do resultado final desta pesquisa.

Ao Professor Fabiano Hartmann Peixoto, que contribui com a tese desde a banca de qualificação, registro meus agradecimentos pelas valorosas sugestões e correções, as quais elevaram o nível da obra.

Também agradeço ao Professor José Renato Gaziero Cella e ao Professor Octavio Campos Fischer, por gentilmente aceitarem integrar a banca de avaliação.

Não poderia deixar de enaltecer os demais Professores que de alguma forma inspiraram este pesquisador durante suas aulas. Assim, agradeço os Professores, Marco Antonio Lima Berberi, William Soares Pugliese, Marcos Augusto Maliska e Ana Claudia Santano por terem contribuído para a pesquisa.

Também agradeço meus pais, Marlene e Antônio Sérgio, e minha avó Dorvalina pelo carinho que sempre me dedicam.

À Fabiana, minha esposa, serei eternamente grato, ante sua compreensão diante dos desafios pessoais e familiares que a pesquisa trouxe.

Saúdo os colegas de PPGD UniBrasil, do doutorado e do mestrado, pelo compartilhamento de ideias e pela mão amiga que estenderam durante o período de doutoramento.

Agradeço à CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), pela bolsa que financiou a pesquisa desde o seu início.

“Existe apenas uma luz na ciência, e acendê-la em qualquer lugar é iluminar todos os lugares” (Isaac Asimov).

RESUMO

A inteligência artificial (IA) tem sido aplicada cada vez mais ao Poder Judiciário brasileiro, demonstrando potencial de grande desenvolvimento. Todavia, o uso de inteligência artificial traz consigo preocupações éticas e jurídicas como, por exemplo: respeito aos Direitos Fundamentais; a independência dos julgadores; erros e vieses de julgamento; e a correta justificação das decisões com o auxílio de IA. Nesse sentido, a tese busca responder se, mediante controle e cooperação recíprocos entre humanos e sistemas inteligentes, existe a formação de um sistema de correção jurídica inerente à interação entre o intérprete humano e a inteligência artificial aplicada às decisões judiciais. Também são explorados aspectos secundários sobre: os potenciais usos da IA judicial; os riscos existentes; os limites dessa interação entre IA e julgador; a natureza dos julgamentos humanos (e seus erros e vieses); em quais aspectos a decisão auxiliada por IA pode ser melhor; e a assimilação do ente humano-máquina em um só ente julgador. A pesquisa é qualitativa e foi desenvolvida mediante os métodos bibliográfico (principal) e exploratório-descritivo (de maneira suplementar), mediante análise interdisciplinar entre Teoria do Direito, Direito Constitucional e Ciência da Computação. O marco teórico principal é a teoria de Robert Alexy, voltada à justificação racional com pretensão de correção, num processo de decisão que inclui diversos tipos de argumentos, o que possibilita o auxílio à pessoa do julgador através de IA. O texto trabalha os conceitos seguintes: robô processual (para atos processuais de maior simplicidade); juiz ciborgue (interação humano-máquina entre o julgador humano e os sistemas de IA); direito fundamental à explicabilidade (prática democrática de exposições das razões adotadas na decisão judicial auxiliada por IA); entropia jurídica (imprevisibilidade informacional prévia das interpretações do Direito); duplo canal de correção (pretensão de correção estabelecida na via comunicacional em julgador humano e sistema de IA judicial, com o aprimoramento argumentativo recíproco); e pré-compromissos democráticos (consensos como parte da racionalidade limitada que pode ser utilizada no desenvolvimento da IA judicial). Os resultados são: sistemas de IA de auxílio à decisão judicial podem ser desenvolvidos de maneira legítima; o controle e a supervisão humanos são necessários; deve ser respeitada a pluralidade epistêmica de cada julgador; o juiz ciborgue surge da interação humano-máquina; os riscos devem ser mapeados, tratados, monitorados e prevenidos; a figura do assessor-desenvolvedor será necessária; impossibilidade de centralização decisória em sistema computacional (independência do julgador); o Direito é também regido pelos processos informacionais (o que leva à entropia jurídica); é necessário um duplo “canal de correção” para reduzir vieses e ruídos (humanos e da IA); a explicabilidade ocorre no limite do tecnicamente possível; há um direito fundamental à explicabilidade; o grau de apoio da IA à decisão judicial pode variar (grau de complexidade e gestão de uso da IA); o Direito deve traduzir o paradigma de *due process of law* para a era da IA; e a argumentação jurídica (justificação interna e justificação externa), com pretensão de correção, é um modelo que garante legitimidade no uso da IA judicial.

Palavras-chave: Inteligência Artificial – Juiz Ciborgue – Duplo Canal de Correção – Poder Judiciário – Argumentação Racional.

ABSTRACT

Artificial intelligence (AI) has been increasingly applied to the Brazilian Judiciary, demonstrating great development potential. However, the use of artificial intelligence causes ethical and legal concerns, such as: respect for Fundamental Rights; the independence of judges; errors and biases in judgment; and the correct justification of decisions with AI support. From this, the thesis answers whether, through reciprocal control and cooperation between humans and intelligent systems, there is the formation of a legal correctness system inherent to the interaction between the human interpreter and the artificial intelligence applied to judicial decisions. Secondary aspects are also explored on: the potential uses of AI in the courts; existing risks; the limits of this interaction between AI and judge; the nature of human judgments (and their errors and biases); in which aspects the AI-supported decision might be better; and the assimilation of the human-machine entity into a single judging entity. The research is qualitative and was developed using bibliographic (main) and exploratory-descriptive (supplementary) methods, through interdisciplinary analysis between Theory of Law, Constitutional Law and Computer Science. The main theoretical framework is Robert Alexy's theory, focused on rational justification with the intention of correctness, in a decision process that includes several types of arguments, which makes AI support possible for the human judge. The text defines the concepts of: procedural robot (for procedural acts of greater simplicity); cyborg judge (human-machine interaction between human judge and AI systems); fundamental right to explainability (democratic practice of expositions of reasons adopted in the judicial decision supported by AI); legal entropy (previous informational unpredictability of interpretations of law); double correction channel (claim to correctness established in the communicational way for human judge to judicial AI system, with the reciprocal argumentative improvement); and democratic precommitments (consensus as part of the bounded rationality that can be used in the development of judicial AI). The results are: AI systems to assist judicial decision-making can legitimately be developed; human control and supervision are necessary; the epistemic plurality of judges must be respected; the cyborg judge emerges from human-machine interaction; the risks must be mapped, treated, monitored and prevented; the figure of the advisor-developer will be necessary; the centralization of the judicial decision in a computer system is not feasible (independence of the judge); the Law is also affected by informational processes (which leads to legal entropy); the double "correction channel" is needed to reduce biases and noise (human and AI); the explainability exists at the limit of what is technically possible; there is a fundamental right to explainability; the degree of AI support for the judicial decision-making may vary (degree of complexity and management of AI use); the Law must be translated the due process paradigm of the AI age; and legal argumentation (internal justification and external justification), with the pretense of correctness, is a model that guarantees legitimacy in the use of judicial AI.

Keywords: Artificial Intelligence – Cyborg Judge – Double Correction Channel – Judiciary Power – Rational Argumentation.

RESUMEN

La inteligencia artificial (IA) se ha aplicado cada vez más en el Poder Judicial de Brasil, demostrando potencial para un gran desarrollo. El uso de la inteligencia artificial trae consigo preocupaciones éticas y jurídicas: el respeto a los Derechos Fundamentales; la independencia de los jueces; errores de juicio y sesgos; y la justificación correcta de las decisiones con la ayuda de la IA. La tesis busca responder si, a través del control recíproco y la cooperación entre humanos y sistemas inteligentes, se configura un sistema de corrección jurídica inherente a la interacción entre el intérprete humano y la inteligencia artificial aplicada a las decisiones judiciales. También se exploran aspectos secundarios: usos potenciales de la IA judicial; riesgos existentes; límites de esta interacción entre AI y juez; naturaleza de los juicios humanos (y sus errores y sesgos); en qué aspectos la decisión asistida por IA puede ser mejor; y la asimilación de la entidad hombre-máquina en una única entidad juzgadora. La investigación es cualitativa y se desarrolló a través de los métodos bibliográfico y exploratorio-descriptivo, a través del análisis interdisciplinario entre Teoría del Derecho, Derecho Constitucional e Ciencia de la Computación. El marco teórico principal es la teoría de Robert Alexy, centrada en la justificación racional con pretensión de corrección, en un proceso de decisión que incluye varios tipos de argumentos, lo que permite ayudar a la persona del juez a través de la IA. El texto trata los siguientes conceptos: robot procedimental (para actos procesales más simples); juez cibernético (interacción hombre-máquina entre el juez humano y los sistemas de IA); derecho fundamental a la explicabilidad (práctica democrática de exponer las razones adoptadas en la decisión judicial asistida por AI); entropía jurídica (imprevisibilidad informativa previa de las interpretaciones de la ley); doble canal de corrección (intención de corrección establecida en la vía de comunicación en juez humano y sistema de IA judicial, con mejora argumentativa recíproca); y precompromisos democráticos (consenso como parte de la racionalidad acotada que puede ser utilizada en el desarrollo de la IA judicial). Los resultados son: los sistemas de IA para ayudar en la toma de decisiones judiciales pueden desarrollarse legítimamente; son necesarios el control y la supervisión humanos; se debe respetar la pluralidad epistémica de cada juez; el juez cibernético surge de la interacción hombre-máquina; los riesgos deben ser mapeados, tratados, monitoreados y prevenidos; será necesaria la figura del asesor-desarrollador; imposibilidad de centralización de la toma de decisiones en un sistema computacional (independencia del juez); el derecho también se rige por procesos informacionales (lo que conduce a la entropía jurídica); se necesita un "canal de corrección" dual para reducir los sesgos y el ruido (humano y de IA); la explicabilidad se produce en el límite de lo técnicamente posible; existe un derecho fundamental a la explicabilidad; el grado de apoyo de la IA para la decisión judicial puede variar (grado de complejidad y gestión del uso de la IA); el derecho debe traducir el paradigma del debido proceso legal a la era de la IA; y la argumentación jurídica (justificación interna y justificación externa), con la pretensión de corrección, es un modelo que garantiza la legitimidad en el uso de la IA judicial.

Palabras clave: Inteligencia Artificial – Juez Cibernético - Doble Canal de Corrección - Poder Judicial - Argumentación Racional.

SUMÁRIO

INTRODUÇÃO	12
1 INTELIGÊNCIA ARTIFICIAL E DECISÃO JUDICIAL: UMA VISÃO DA TEORIA DO DIREITO SOBRE OS ELEMENTOS PARA A LEGITIMIDADE DA IA JUDICIAL	23
1.1 Transcendência Processual: A Constante Evolução Tecnológica	23
1.2 Inteligência Artificial: Aspectos de Um Conceito Não Jurídico	31
1.3 Alvorecer: Inteligência Artificial e Direito	34
1.4 O Juiz Ciborgue: Riscos e Possibilidades da Interação entre Inteligência Artificial e Teoria da Decisão Judicial	38
1.5 Entropia e Direito: P vs. NP, ou, o Mundo é Maior que o Direito	56
1.6 Da Entropia Jurídica ao Duplo Canal de Correção: Direito e Teoria da Informação	59
2 PRINCIPAIS ASPECTOS DA EXPLICABILIDADE DA INTELIGÊNCIA ARTIFICIAL APLICADA AO DIREITO	89
2.1 Inteligência Artificial Judicial: A Construção da Ideia de Explicabilidade	89
2.2 Caminhos do Direito à Explicabilidade: A Objetividade Possível para a Explicabilidade de Decisões Judiciais Auxiliadas pela IA	99
2.3 IA Judicial Explicável: Direito à Explicabilidade e Deliberação Racional	114
2.4 Deliberação Racional e IA Judicial: a Explicabilidade como Elemento da Justiça	118
2.5 <i>Due xAI of Law</i> : o Direito Fundamental à Explicabilidade	127
2.6 Inteligência Artificial e Transparência <i>by Design</i> : a Explicabilidade Necessária e a Interpretabilidade Possível	134
2.7 IA Explicável (xAI) e Transparência Possível: Os limites da Explicabilidade da Inteligência Artificial nas Decisões Judiciais	143
2.8 Aspectos Dogmáticos do Direito à Explicabilidade	153
3 UMA COOPERAÇÃO POSSÍVEL ENTRE INTELIGÊNCIA ARTIFICIAL E DECISÃO JUDICIAL	166
3.1 Aspectos da Objetividade dos Sistemas da IA Judicial	166

3.2 Robô Processual, Regra Padrão e Esquecimento: Heurísticas nos Atos Processuais de Maior Simplicidade	172
3.3 Inteligência Artificial Judicial: a Pluralidade Epistêmica e o Juiz da Era da IA	182
3.4 Pré-Compromissos, Heurísticas e Inteligência Artificial Judicial	195
3.4.1 A Caracterização dos Consensos como Heurísticas para a IA Judicial	195
3.4.2 Heurísticas e Pré-Compromissos: Heurísticas do Julgador <i>versus</i> Heurísticas do Direito	206
 4 JUSTIFICAÇÃO DA DECISÃO JUDICIAL E INTELIGÊNCIA ARTIFICIAL	 220
4.1 O Juiz Ciborgue e a Pretensão de Correção: Facetas da Interação entre Suporte Fático, Fatores (“<i>Factors</i>”) e Conceitos Jurídicos Intermediados	220
4.1.1 O Suporte Fático Percebido pela Inteligência Artificial	220
4.1.2 Fatores (“ <i>factors</i> ”) e a Representação do Suporte Fático	227
4.1.3 Conceitos Jurídicos Intermediados (“ <i>ILC’s</i> ”), o Suporte Fático Hipotético e a Abordagem Híbrida	234
4.2 A Justificação Interna das Decisões Tomadas com Auxílio de Inteligência Artificial	237
4.3 Possibilidades da Justificação Externa na Inteligência Artificial Judicial	252
4.3.1 Justificação Externa e IA Judicial: da Pretensão de Correção Normativa à Pretensão de Correção Moral	252
4.3.2 Elementos da Justificação Externa: Regras do Discurso Racional e sua Relação com a IA Judicial	268
4.4 IA e Razoabilidade: do Contexto Fático ao Teste Lógico nas Proposições Normativas Auxiliadas por Inteligência Artificial	287
4.4.1 O Razoável: Os Limites Fáticos e a Expressão da Razoabilidade através da IA	287
4.4.2 Razoabilidade e IA: Do Teste <i>Wednesbury</i> ao Teste de Idoneidade Lógica	296
 5 ASPECTOS DO DESENVOLVIMENTO DE INTELIGÊNCIA ARTIFICIAL JUDICIAL: A INFLUÊNCIA DA PRETENSÃO DE CORREÇÃO EM SISTEMAS DE IA JURÍDICA	 304
5.1 Aspectos Gerais da IA Judicial: Uma Visão Para Além dos Robôs	305

5.2 A Obtenção das Bases Dados (<i>Datasets</i>) para o Desenvolvimento de IA Judicial	310
5.3 <i>Data Science</i> e <i>Machine Learning</i> : Uma Visão Necessária do Aprendizado de Máquina voltado à IA Judicial	312
5.4 Ciência de Dados e Direito: Principais Aplicações de Modelos de IA Judicial Baseados em Aprendizado de Máquina	315
5.5 <i>Machine Learning</i> : Aprendizagem Supervisionada, Aprendizagem Não Supervisionada e Aprendizagem por Reforço	323
5.6 A Pretensão de Correção no Ciclo da Ciência de Dados Jurídica: O Desenvolvimento de Modelos de <i>Machine Learning</i> para a IA Judicial	327
5.7 <i>Deep Learning</i> : O Aprendizado Profundo e sua Aplicabilidade para o Juiz Ciborgue	347
5.8 Processamento de Linguagem Natural: Uma Conversa entre a IA e o Juiz Ciborgue	353
CONCLUSÃO	361
REFERÊNCIAS	366

INTRODUÇÃO

Nas peças de teatro da Grécia antiga, a fim de garantir uma solução rápida para situações existentes na narrativa se inseria, por uma espécie de maquinário, a figura metafórica de um deus no palco. Assim, convencionou-se denominar tal solução retórica de deus *ex machina*.¹

Desde a antiguidade o maquinário inventado pela humanidade passou de rodas, polias e cordas para a grande tecnologia do microchip; das máquinas de calcular avançadas no início do século XX (como os computadores utilizados para calcular trajetórias balísticas e para decifrar a criptografia do código nazista *Enigma*), placa de silício, dos processadores e dos *softwares*, o que possibilita uma realidade hoje bastante consistente na aplicação de Inteligência Artificial em variados campos do conhecimento.²

Nesse sentido, Isaac Asimov foi visionário em sua obra *Eu, Robô* ao retratar máquinas que se imaginavam vivas porque conseguiam “pensar” e tomar decisões alegadamente racionais, ou muitas vezes mais irracionais e equivocadas do que um ser humano.³ A esse respeito, se tornou célebre a propositura por Asimov das três leis da robótica, com destaque para a segunda lei da robótica que prevê que um robô deve obedecer às ordens dadas por seres humanos, exceto nos casos em que tais ordens possam causar “mal” a um ser humano.⁴

Na medida em que o futuro antevisto por Asimov parece cada vez mais próximo, muito pelo aumento exponencial da capacidade de processamento dos

¹ “*Bearing in mind classical Greek images of the winged goddess Nike hovering over victors’ heads, the royal engineers created an immense statue of the goddess, suspended on cables out of sight. Similar deus ex machina technology was used on the stage in classical Greek theatrical performances, but this scheme was off the scale*” (MAYOR, Adrienne. **Gods and Robots: Myths, Machines and Ancient Dreams of Technology**. Princeton: Princeton University Press, 2018, posição 3211-3213. E-book Kindle).

² KURZWEIL, Ray. **Como Criar uma Mente: os segredos do pensamento humano**. São Paulo: Aleph, 2014, posição 171-175. E-book Kindle.

³ “– Passei os últimos dois dias em concentrada introspecção – disse Cutie [o robô] – e os resultados foram muito interessantes. Comecei com a única suposição segura que senti que me era permitido fazer. Eu existo porque eu penso...” (...)

“– Você está sendo ridículo. Eu já lhe disse que nós o fizemos.

– E se não acredita em nós – acrescentou Donovan –, ficaremos felizes em desmontá-lo!

O robô fez um gesto largo com as mãos, em desaprovação.

– Não aceito nada com base em autoridades no assunto. Uma hipótese deve ser sustentada pela razão, caso contrário não tem valor... e supor que vocês me fizeram vai contra todos os preceitos da lógica” (ASIMOV, Isaac. **Eu, Robô**. Tradução: Aline Storto Pereira. São Paulo: Aleph, 2015, posição 922-931. E-book Kindle).

⁴ Ibid., posição 687-688.

computadores⁵, o ser humano se concentra no desenvolvimento tecnológico da Inteligência Artificial para a solução de problemas e para a melhoria da vida em sociedade em diversas áreas, como saúde, engenharia e programação, por exemplo.⁶

Por outro lado, um pouco do contexto social sobre o uso de algoritmos de Inteligência Artificial pode observado na música *Algorithm* da banda britânica *Muse*, que sintetiza o sentimento de parte da humanidade quando confrontada com a tecnologia:

Queime como um escravo
Mova como uma engrenagem
Estamos enjaulados em simulações
Algoritmos evoluem
Empurram-nos de lado e nos tornam obsoletos.⁷

Noutra visão, com um ar mais desprendido existem os versos da canção *I Love My Computer* da banda *Bad Religion*:

Tudo o que eu preciso fazer é clicar em você
E vamos ficar juntos da forma mais desalmada
E nunca vamos arruinar o dia um do outro
Pois quando me canso, eu só clico
E você vai embora.⁸

⁵ MOORE, Gordon E. Cramming more components onto integrated circuits. In: **Electronics**, vol. 38, n. 8, April 19, 1965.

⁶ FERRUCCI, David et al. Watson: Beyond Jeopardy! In: **Artificial Intelligence**, 199-200, 2013, p. 93-105.

⁷ Tradução livre de:

*"Burn like a slave
Churn like a cog
We are caged in simulations
Algorithms evolve
Push us aside and render us obsolete".*

(BELLAMY, Matthew. *Algorithm*. In: MUSE. **Simulation Theory**. Londres: Warner Bros. Records e Helium-3, 2018. Disponível em: <https://youtu.be/wJcVLxRtDms>. Acesso em: 01 jun. 2022).

⁸ Tradução livre de:

*"All I need to do is click on you
And we'll be joined the most soulless way
And we'll never ruin each other's day
'Cause when I'm through I just click
And you just go away".*

(GRAFFIN, Greg. *I Love My Computer*. In: BAD Religion. **The New America**. Nova York: Atlantic Records, 2000. Disponível em: <https://youtu.be/zlM0vahvauU>. Acesso em: 01 jun. 2022).

Ainda que não seja algo novo na sociedade, vem ganhando densidade o sentimento antagônico de, por um lado, ser substituído pelos algoritmos e computadores e, por outro, ser simplesmente auxiliado pela tecnologia. Tal sentimento inundou o espaço jurídico e vem causando um estranhamento entorpecente até mesmo entre os estudiosos do Direito.

Desse modo, falar de Inteligência Artificial e seus impactos jurídicos e filosóficos é caminhar pela sala dos desejos da humanidade. Os sonhos de criação de inteligências artificiais amistosas à humanidade e os pesadelos sobre uma IA maléfica habitam o imaginário cotidiano. Conviver com algo próximo da figura amistosa de *C-3PO* de *Star Wars* pode parecer mais próximo do que conhecer a maquinaria desumanizante de tantas outras obras de ficção.⁹ Para isso o Direito deve se debruçar sobre as diversas questões que os avanços da Inteligência Artificial lhe causam.

O desenvolvimento de sistemas de Inteligência Artificial com a finalidade de auxílio à atividade jurisdicional, para a realização de atos de impulso do processo ou para o auxílio à decisão judicial, vem crescendo exponencialmente. Essa aceleração da aplicação da Inteligência Artificial ao Direito, em especial ao Poder Judiciário, é uma realidade fruto de um histórico de implementação de tecnologias, com destaque para o processo judicial eletrônico e, na sequência, para a digitalização do acervo de processos na maioria dos Tribunais, o que criou infraestrutura de tecnologia da informação e base de dados sólida que servem como o trampolim para o próximo estágio de desenvolvimento pelo qual se está passando através da adoção da IA jurisdicional.

É relativamente recente no Brasil a inclusão de inteligência artificial nas rotinas jurídicas, especialmente com a finalidade de apoio ao Poder Judiciário. Todavia, salta aos olhos o crescimento exponencial do número de projetos de aplicação de Inteligência Artificial ao Judiciário. Segundo pesquisas da Fundação Getúlio Vargas, no ano de 2022 já existiam 64 projetos de Inteligência Artificial judicial em andamento¹⁰. Por outro lado, foram catalogados 97 projetos de IA no judiciário

⁹ BOSTROM, Nick. **Superinteligência**: Caminhos, perigos, estratégias. Tradução: Aurélio Antônio Monteiro, Clemente Gentil Penna, Fabiano Geremias Monteiro e Patrícia Ramos Geremias. Rio de Janeiro: Darkside Books, 2014, posição 8864-8877. E-book Kindle.

¹⁰ SALOMÃO, Luis Felipe (coord.) et al. **Inteligência Artificial**: Tecnologia Aplicada à Gestão dos Conflitos no Âmbito do Poder Judiciário Brasileiro. 2ª ed. FGV Conhecimento: Centro de Inovação, Administração e Pesquisa do Judiciário, 2022. Disponível em: https://ciapj.fgv.br/sites/ciapj.fgv.br/files/relatorio_ia_2fase.pdf. Acesso em: 01 jun. 2022.

brasileiro.¹¹ No ano de 2020, os diversos tribunais brasileiros relataram a existência 72 projetos em andamento para a aplicação de Inteligência Artificial pelo Poder Judiciário.¹² O maior exemplo se trata do projeto *Victor* do Supremo Tribunal Federal.¹³

Dessa forma, a inovação decorrente das abordagens possíveis através da Inteligência Artificial traz consigo diversas preocupações éticas e jurídicas que se mostram naturais, como, por exemplo, o respeito aos Direitos Fundamentais, a aferição da independência dos julgadores, os possíveis erros e vieses de julgamento e a adequação da justificação das decisões e a correção da aplicação do Direito com o auxílio de sistemas de Inteligência Artificial.

Neste momento já se pode vislumbrar as primeiras facetas daquilo que tem se convencionado chamar de revolução 4.0¹⁴, um contexto fático de relevância histórica, econômica e social que cada vez mais hibridiza máquinas e humanidade num só corpo. Por outro lado, também surgem abordagens preocupadas com a utilização da Inteligência Artificial numa armadilha voltada à instrumentalização do comportamento humano para fins de “modificação, predição, monetização e controle”.¹⁵

A era da Inteligência Artificial só está começando. Isso traz a reflexão sobre quais características hoje adjetivadas como “humanas” restarão num mundo de uma sociedade mergulhada em tecnologia.

Observar a realidade para interpretá-la, interpretar a realidade para se posicionar, justificar o raciocínio para decidir, decidir racionalmente para julgar de forma legítima. Eis os mantras que transcendem a teoria do direito, o direito constitucional e o direito processual através de diferentes facetas, bem como que têm inspirado intensos debates na academia, no Judiciário e na sociedade.

¹¹ RAMOS, Janine Vilas Boas Gonçalves. **Inteligência Artificial no Poder Judiciário Brasileiro: Projetos de IA nos tribunais e o sistema de apoio ao processo decisório judicial**. São Paulo: Dialética, 2022, p. 267. E-book Kindle.

¹² JOTA. **Judiciário brasileiro tem ao menos 72 projetos de inteligência artificial nos tribunais**: Ferramentas são usadas para auxiliar agrupamento de demandas repetitivas a até para sugerir minutas. [2020]. Disponível em: <https://www.jota.info/coberturas-especiais/nova-e-acao/judiciario-brasileiro-tem-ao-menos-72-projetos-de-inteligencia-artificial-nos-tribunais-09072020>. Acesso em: 20 ago. 2020.

¹³ HARTMANN PEIXOTO, Fabiano. Projeto Victor: Relato do Desenvolvimento da Inteligência Artificial na Repercussão Geral do Supremo Tribunal Federal. **Revista Brasileira de Inteligência Artificial e Direito**, v. 1, n. 1, jan-abr. 2020.

¹⁴ SCHWAB, Klaus. **A Quarta Revolução Industrial**. Tradução: Daniel Moreira Miranda. São Paulo: Edipro, 2018, posição 2425. E-book Kindle.

¹⁵ ZUBOFF, Shoshana. **A Era do Capitalismo de Vigilância**: A luta por um futuro humano na nova fronteira de poder. Tradução: George Schlesinger. Intrínseca: Rio de Janeiro, 2019, p. 521. E-book Kindle.

A abordagem padrão do problema tem sido discutir as questões inerentes à teoria da norma jurídica, à interpretação do direito, à argumentação justificatória e aos elementos objetivos da decisão judicial. Tudo isso faz parte do legado teórico que trouxe a democracia até o momento presente. Todavia, a abordagem tradicional sobre a decisão judicial (também lembrada em certos aspectos neste estudo) apresenta-se como uma heurística em si mesma, a qual merece maiores detalhamentos para se desvelar aspectos da teoria da decisão judicial.

A partir disso, o tema da legitimidade e dos limites da inclusão de sistemas de Inteligência Artificial na decisão judicial ganha destaque. Todavia, percebe-se que o enfrentamento da questão apenas através de um dos fenômenos não traz respostas que possam autorizar uma conclusão serena sobre o tema, para além das caricaturas da utilização da IA no Direito.

Em relação à hipótese, a presente tese busca responder **se no processo de decisão judicial, mediante controle e cooperação recíprocos entre humanos e sistemas inteligentes, existe a formação de um sistema de correção jurídica inerente à interação entre o intérprete humano e a Inteligência Artificial**. Por se tratar de um tema de relativa novidade científica, explorado ainda de maneira incipiente pela doutrina, algumas perguntas secundárias são exploradas no decorrer do texto, em específico a respeito dos potenciais usos da IA judicial, dos riscos existentes, dos limites dessa interação entre IA e julgamento e da própria natureza dos julgamentos. A esse respeito a decisão exclusivamente humana é livre de erros e vieses? A decisão maquinal através das IA pode ser melhor em quais aspectos? Há a assimilação do ente humano-máquina em um só ente julgador?

Assim, esta pesquisa é de cunho qualitativo e foi desenvolvida mediante os métodos bibliográfico (principal) e exploratório-descritivo (de maneira suplementar), por meio de análise interdisciplinar entre Teoria do Direito, Direito Constitucional e Ciência da Computação. Dessa forma, a pesquisa levou em conta institutos e conceitos da Teoria do Direito, do Direito Constitucional e da Ciência da Computação, todavia também direcionou a análise para a exploração prática da programação (Python, principalmente), de Ciência de Dados (mediante análise de modelos de aprendizado de máquina) e de sistemas de IA já existentes (*GPT-3* e *CARA AI*, por exemplo), a fim de que a experiência e a teoria pudessem levar a análise da interação entre IA e Direito para um plano concreto e maturado a respeito do problema. Dessa

forma, embora a abordagem metodológica tenha sido a partir do Direito (e o autor teve o seu ponto de partida), a experimentação prática na Ciência da Computação teve papel importante na construção da tese.

Para a presente tese, o marco teórico principal é a teoria do direito de Robert Alexy, escolhida por expressar a possibilidade de justificação racional através de um procedimento de correção relativa que engloba aspectos normativos e morais, mediante razão prática e um processo de decisão que pode incluir a participação de diversos tipos de argumentos, podendo se abrir para novas possibilidades de auxílio à pessoa do julgador através de IA, uma vez que o julgador é o responsável pela decisão no processo argumentativo. Nesse aspecto, a necessidade de justificação racional é auxiliada pela concepção democrática inerente ao uso público da razão (Jürgen Habermas) e à explicabilidade de modelos de Inteligência Artificial (trabalhada a partir de autores de referência na Ciência da Computação). Dessa maneira, ainda em relação ao marco teórico principal, é trazido o pensamento de Eugenio Bulygin sobre a concepção de sistema jurídico como um microssistema fechado logicamente em si mesmo, a fim de se possibilitar a construção de soluções através de IA judicial com pretensão de consistência, relativa independência e que, dentro de seu escopo (domínio), busquem completude (ainda que esta não seja alcançada ou deva ser desafiada pela justificação do julgador humano). Nesse caminho, a análise da busca de objetividade e racionalidade do julgamento auxiliados por meio de IA judicial passa por leituras auxiliares de autores como: Claude Shannon e sua entropia, como parte do processo informacional, que é tratada através de canais de correção; Ilya Prigogine e sua abordagem sobre o reconhecimento da incerteza inerente à entropia do Universo; Daniel Kahneman, Richard Thaler e Cass Sunstein e seus estudos sobre as heurísticas, regras padrão e vieses de julgamentos humanos; Neil MacCormick e a sua argumentação jurídica pautada numa visão institucional e consequencialista; Herbert Hart, Hans Kelsen e Joseph Raz na busca de delinear os elementos de objetividade do ordenamento jurídico que influenciam o processo de decisão; o papel do princípio da sinceridade na teoria do agir comunicativo de Jürgen Habermas mediante sua influência sobre a teoria de Robert Alexy, o que por sua vez repercute na presente pesquisa; e os estudos mais influentes sobre a relação entre Inteligência Artificial e Direito existentes no Brasil (com destaque para o trabalho de Fabiano Hartmann Peixoto).

Na presente obra, deve ser destacado que os autores referidos conversam em níveis diferentes, seja pelas fontes de seus pensamentos ou pelas características das questões por eles originalmente tratadas. Todavia, os diferentes níveis e as peculiares abordagens de cada um dos autores podem ser aproveitados, dentro do contexto da tese, para explicar o entrelaçamento entre direito e Inteligência Artificial, sem que sejam desconsiderados os diferentes pontos de partida e os objetivos das obras de cada um dos autores. Desse modo, o caráter multidisciplinar da pesquisa exigiu um olhar que englobasse diferentes pontos de vista, o que se buscou fazer com respeito ao pensamento de cada pesquisador (jurista ou cientista) que, de alguma forma, foi parte da construção desta tese.

Além disso, de plano deve ser observado que a tese não trata de temas paralelos, como, por exemplo, a exclusão digital de grande parte da população brasileira e os efeitos disso sobre o modelo de Justiça 4.0, pois este não é escopo do presente estudo, que é focado no processo de julgamento auxiliado por sistemas de Inteligência Artificial. Feito este aviso necessário, passa-se a descrição dos capítulos da presente obra.

O Capítulo 1 trata das principais abordagens da Teoria do Direito na busca de investigar a legitimidade da inclusão de sistemas de Inteligência Artificial no processo de decisão judicial. Dessa forma, foca na análise da evolução da tecnologia utilizada nos processos judiciais; na definição de um conceito amplo de IA (conforme adotado na Ciência da Computação); no estudo dos efeitos (positivos e negativos) da IA sobre a independência dos julgamentos (como os vieses do julgador humano a serem alertados pela IA e os vieses inoculados na IA a serem suprimidos pelo juiz) e do reconhecimento da entropia informacional como um fenômeno que alcança o Direito, que pode ser objeto de um duplo canal de correção (Claude Shannon) na relação Juiz-IA (pensamento humano-máquina¹⁶), no qual a Inteligência Artificial judicial se manifesta como uma via comunicacional de estímulo argumentativo ao julgador humano (e este em relação à IA, reciprocamente), o que se constitui no modelo de juiz ciborgue.¹⁷

¹⁶ “Há duas maneiras de abordar o problema: pode-se considerar as máquinas como parasitas que se enxertam na atividade orgânica e dela tiram benefícios, ou os humanos como parasitas das máquinas tirando proveito de sua existência. A vantagem sendo mútua e parasitismo sendo recíproco, devemos falar de simbiose” (HOQUET, Thierry. **Filosofia Ciborgue**: pensar contra os dualismos. Tradução: Marcio Honório de Godoy. São Paulo: Perspectiva, 2019, p. 86).

¹⁷ Cunhado originalmente no contexto da corrida espacial como um acrônimo para “*cybernetic organism*”, o termo “cyborg” (ciborgue) leva em conta a incorporação deliberada de componentes

Na sequência, o Capítulo 2 trabalha a ideia de explicabilidade de decisões tomadas com auxílio de Inteligência Artificial, tendo por base as discussões, inerentes à Ciência da Computação, relativas ao esforço para construção de sistemas de

externos como forma de ampliar a função de controle autorregulatório do organismo, visando possibilitar sua adaptação a novos ambientes (CLYNES, Manfred E.; KLINE, Nathan S. *Cyborgs and Space*. **Astronautics**, p. 29-33, Sep. 1960).

Já primeira referência conhecida ao termo “cyborg judge” (“juiz ciborgue”) remonta ao seguinte artigo: DUNAGAN, Jake. *Neuro-Futures: The Brain, Politics, and Power*. **Journal of Futures Studies**, Taipei, v. 9, n. 2, p. 1-18, 2004.

No referido artigo o autor Jake Dunagan conceitua a ideia de “subjetividade ciborgue” (“cyborg subjectivity”) como inerente aos “aumentos neurais” trazidos por tecnologias disruptivas capazes de produzir “novas subjetividades” através de um aprimoramento cerebral, o que remeteria, segundo o autor, aos conceitos de ciborgue e pós-humano, de maneira que a tecnologia representaria um domínio sobre o “eu”. Assim, ao lembrar de precedente já superado da Suprema Corte dos Estados Unidos [Buck v. Bell. 274 U.S. 200 (1927)], o autor questiona se um futuro “juiz ciborgue” (“cyborg judge”) obrigaria as pessoas ao aprimoramento cognitivo com o uso de tecnologias cibernéticas como espécie de política baseada em eugenia (Ibid., p. 7).

Por outro lado, destaca-se que a presente pesquisa adota postura conciliatória entre o “eu” humano e a tecnologia da inteligência artificial, conforme será explicitado no decorrer do texto, o que difere da posição de Dunagan, ainda que se reconheça a relevância de seu pensamento.

Ademais, o uso do termo “cyborg judge” no texto de Dunagan foi feito de passagem e num contexto filosófico mais amplo, sem preocupação com a delineação dos elementos de caracterização de um conceito jurídico sobre o tema.

Em 2009, Raquel Prata publicou um artigo (PRATA, Raquel. O Juiz Cyborg - A Informática e o Direito. **Revista Julgar**, Lisboa, n. 8, 2009) em que utilizou o termo “Juiz Cyborg”. Contudo, o texto limita-se a criticar genericamente a utilização de tecnologia informática pelo Judiciário, deixando de definir uma conceituação sobre o tema do juiz ciborgue (que se limitou à chamada de título) ao não tratar de qualquer questão realacionada à automatização ou à inteligência artificial.

Já em novembro de 2019, Rebecca Crotoft publicou artigo (CROTOFT, Rebecca. 'Cyborg Justice' and the Risk of Technological-Legal Lock-In. **Columbia Law Review Forum**, New York, v. 119, p. 233-251, 20 nov. 2019) em que, mediante uma exploração inicial da temática, reconhece a possibilidade de colaboração entre juizes humanos e sistemas de Inteligência Artificial. Assim, a referida jurista pode ser considerada a autora do termo “*Cyborg Justice*” (“Justiça Ciborgue”), embora não tenha delineado os elementos definidores de um “juiz ciborgue”.

Nesse sentido, o termo “juiz ciborgue” e sua tradução “cyborg judge”, quando utilizados pelo presente autor, remetem ao primeiro artigo jurídico escrito cujo título o empregou com sentido que a tese desenvolve: PÁDUA, Sérgio Rodrigo de. O Juiz Ciborgue: Inteligência Artificial e Decisão Judicial. In: ROVER, Aires José; et al. (coord.). **Direito, Governança e Novas Tecnologias III**. Florianópolis: CONPEDI, 2020, p. 315-335.

Referido artigo foi escrito no primeiro semestre de 2019 e publicado no ano de 2020, após sua apresentação no II Encontro Virtual do CONPEDI (Conselho Nacional de Pesquisa e Pós-Graduação em Direito), texto que fez parte do caminho de construção da presente pesquisa. Este artigo traçou as primeiras bases da abordagem deste autor para a ideia de cooperação entre inteligência artificial e pensamento humano, que ocorreu na busca do aprimoramento da decisão judicial pautado no respeito ao “eu” do julgador e à sua legitimidade, o que se fez evocando-se o conceito de um ciborgue, um ser hibridizado parte cibernético e parte orgânico. Em suma, na linha adotada nesta obra, é marcante a diferença entre o conceito de “juiz ciborgue” deste autor e a referência feita por Jake Dunagan, bem como há um grande aprofundamento na problemática do juiz ciborgue, o que não se observa no texto de Rebecca Crotoft.

Além disso, percebe-se que outra pesquisa jurídica que cita o termo “cyborg judge”, tradução do termo romeno “judecătorul cyborg”, trata-se de artigo que foi publicado no ano de 2022: SPOREA, Daniel-Cosmin. Judecătorul Cyborg: O Automatizare a (In)Justiției Penale? **Revista Universul Juridic**, București, n.1, p. 71-79, 2022.

Logo, levando-se em conta as bases de pesquisa conhecidas, o conceito estritamente jurídico “juiz ciborgue” é de autoria deste autor, sendo que as suas menções durante o presente texto deixarão de ser repetidamente referenciadas.

Inteligência Artificial explicável (xAI). Todavia, passando a noção de um direito fundamental à explicabilidade da decisão judicial auxiliada pela IA, que se lastreia na concepção do uso público da razão (Jürgen Habermas) e na justificação com pretensão de correção (Robert Alexy) para se criar sistemas de IA judicial. Nesse capítulo, a abordagem de necessária transparência *by design*, bem como o reconhecimento das limitações técnicas à explicabilidade completa (*full explainability*), levam a adoção de uma distinção em explicabilidade (em sentido estrito), como a possibilidade da compreensão humana dos motivos pelos quais um modelo de IA Judicial sugeriu determinada solução, e interpretabilidade, consistente na compreensão técnica do funcionamento de algoritmos e modelos de IA (o que depende de conhecimento em Ciência da Computação). O referido capítulo também trata dos aspectos dogmáticos do Direito à Explicabilidade.

Já o Capítulo 3 buscará tratar da possibilidade de utilização de heurísticas do julgador para a implementação de regras padrão para a prática de atos processuais de maior simplicidade; a necessidade de respeito à pluralidade epistêmica acerca da Teoria do Direito, que embasa as decisões de cada juiz como um elemento da democracia, da independência judicial e do pluralismo que deverá pautar o desenvolvimento de sistemas de IA judicial. Além disso, trata dos pré-compromissos democráticos como espécie de consenso constitucional a pautar as heurísticas do julgador que serão integradas nos sistemas de auxílio à decisão judicial através de Inteligência Artificial.

O Capítulo 4 aprofunda a interação entre a Inteligência Artificial e o processo de justificação da decisão judicial através da argumentação jurídica, passando pelas fases de justificação interna e justificação externa da decisão, bem como os cuidados necessários com cada tipo de argumento a ser construído com o auxílio de Inteligência Artificial, a fim possibilitar a programação ou a modelagem de sistemas de IA judicial que tenham eficiência computacional e que garantam o grau de correção adequado. Dessa forma, o Capítulo 4 trata do desafio da computação da facticidade do Direito na busca de garantir sugestões de aplicação do Direito pela IA que sejam lógicas para cada julgamento, o que passa por uma visão do suporte fático do direito aplicado e das consequências das decisões.

O Capítulo 5 faz uma análise interdisciplinar, mediante um estudo contextualizado da Ciência de Dados e da Ciência da Computação aplicáveis à

decisão judicial, a fim de demonstrar aspectos técnicos do processo de desenvolvimento da IA aplicada ao Direito, de maneira a relacionar isso com as categorias trabalhadas ao longo dos demais capítulos, especialmente para que exista a percepção de como o desenvolvimento de sistemas de Inteligência Artificial judicial pode contribuir para a prática de decisões judiciais que levem em conta a pretensão de correção.

Dessa forma, o desafio que se colocou à presente tese é a ressignificação (ainda que parcial) da Teoria do Direito, na medida em que as bases do *mainstream* do pensamento jurídico são colocadas sob desafios que, até pouco tempo, eram vistos como simples ficção científica ou, no mínimo, algo longínquo e de pouca importância prática. Outro desafio é o senso comum teórico¹⁸ (ao menos no Direito) de que o ser humano é um ser sempre racional, embora esse esteja preso ao próprio efeito halo de que o humano é a medida de todas as coisas.

Sem a necessidade de uma abordagem estatística, o ser humano (e a coletividade dos humanos, a humanidade) decide racionalmente só em termos justificatórios, pois a razão lógico-analítica é um emaranhado de pensamentos e gatilhos mentais muito bem maquiados pelas palavras e ações humanas.

Assim, o ser humano é falível e suas teorias são falseáveis.¹⁹ O reconhecimento das limitações humanas é um passo humildemente necessário para se admitir que a forma de decisão humana sempre pode ser aprimorada. Essa concepção inspirou a presente pesquisa.

Por outro lado, a presente tese leva em conta que computadores e a Inteligência Artificial não são remédio para todas as dores, porém podem trazer mais humanidade integrada neles do que em muitas mentes humanas.

¹⁸ “Os processos significativos do direito apresentam-se como um conjunto heterogêneo de hábitos semiológicos de referência (senso comum teórico) e de discursos organizados a partir de ditos hábitos. Ou seja: os primeiros operam como um código para as enunciações jurídicas.

Metaforicamente, caracterizamos o senso comum teórico como a voz ‘off’ do direito, como uma caravana de ecos legitimadores de um conjunto de crenças, a partir das quais, podemos dispensar o aprofundamento das condições e das relações que tais crenças mitificam” (WARAT, Luís Alberto. *Saber Crítico e Senso Comum Teórico dos Juristas*. **Seqüência**: Estudos Jurídicos e Políticos, Florianópolis, v. 3, n. 5, p. 48-57, 1982. DOI: 10.5007/%25x).

¹⁹ “O critério de demarcação inerente à Lógica Indutiva – isto é, o dogma positivista do significado – equivale ao resqúcio de todos os enunciados da ciência empírica (ou todos os enunciados ‘significativos’) devem ser suscetíveis de serem, afinal, julgados com respeito à sua verdade e falsidade; diremos eles devem ser ‘conclusivamente julgáveis’. Isso quer dizer que sua forma deve ser tal que se torne logicamente possível verificá-los e falsificá-los” (POPPER, Karl. **A Lógica da Pesquisa Científica**. Tradução: Leonidas Hegenberg e Octanny Silveira da Mota. 2ª ed. São Paulo: Cultrix, 2013, p. 37-38).

Cabe destacar que durante os anos de 2019 a 2023 esta pesquisa nasceu sob a realidade que se impôs. No início existiam poucos acadêmicos brasileiros que tratavam a respeito da possibilidade do uso da IA no Direito. Muitos textos foram publicados no período, mas nenhum que tenha tratado do núcleo da presente tese, o que denota que esta pesquisa ganhou corpo ao se perceber que se agigantaram as discussões sobre o tema da Inteligência Artificial jurídica.

A beca e toga sempre terão seu lugar de destaque, mas agora estão imersas no oceano de algoritmos que passa a fazer parte do mundo jurídico.

1 INTELIGÊNCIA ARTIFICIAL E DECISÃO JUDICIAL: UMA VISÃO DA TEORIA DO DIREITO SOBRE OS ELEMENTOS PARA A LEGITIMIDADE DA IA JUDICIAL

1.1 Transcendência Processual: A Constante Evolução Tecnológica

Se, no início, tudo era verbo, agora tudo aparenta ser compreendido com base nos algoritmos. Todavia, o salto para a era dos algoritmos (na qual o mundo do direito ainda está “molhando os pés”) somente foi possibilitado por uma infinidade de avanços tecnológicos. Assim, se atualmente são estudadas as possibilidades de aplicação de Inteligência Artificial ao direito, há que se destacar a evolução da tecnologia até o momento presente.

Desde a antiguidade, o ser humano tem sonhado com entidades que lhes auxiliem a fazer seu trabalho²⁰, entidades aptas a aprimorar as habilidades humanas e a esculpir a partir da pedra bruta novas visões de mundo, renovados os sentimentos do devir cultural inerente à humanidade que todo *Homo sapiens* carrega consigo.²¹

Dessa visão, percebe-se que o próprio Estado foi uma entidade maquinal criada como maneira eficiente de organizar a sociedade. A burocracia estatal, como ressalta Max Weber²², é a síntese da racionalidade no alinhamento dominante de engrenagens humanas e institucionais na busca de uma finalidade coletiva.

²⁰ “With a few exceptions, in the myths as they have survived from antiquity, the inner workings and power sources of automata are not described but left to our imagination. In effect, this nontransparency renders the divinely crafted contrivances analogous to what we call “black box” technology, machines whose interior workings are mysterious. Arthur C. Clarke’s famous dictum comes to mind: the more advanced the technology, the more it seems like magic. Ironically, in modern technoculture, most people are at a loss to explain how the appliances of their daily life, from smartphones and laptops to automobiles, actually work, not to mention nuclear submarines or rockets. We know these are manufactured artifacts, designed by ingenious inventors and assembled in factories, but they might as well be magic. It is often remarked that human intelligence itself is a kind of black box. And we are now entering a new level of pervasive black box technology: machine learning soon will allow Artificial Intelligence entities to amass, select, and interpret massive sets of data to make decisions and act on their own, with no human oversight or understanding of the processes. Not only will the users of AI be in the dark, but even the makers will be ignorant of the secret workings of their own creations. In a way, we will come full circle to the earliest myths about awesome, inscrutable artificial life and biotechné” (MAYOR, Adrienne. Op. cit., posição 255-265).

²¹ HARARI, Yuval Noah. **Sapiens**: Uma breve história da humanidade. Porto Alegre: L&PM Editores, 2015, posição 148-152. E-book Kindle.

²² WEBER, Max. **Economia e Sociedade**: Fundamentos da sociologia compreensiva. Vol. 2. Tradução: Regis Barbosa e Karen Elsabe Barbosa. Brasília: Editora Universidade de Brasília, 1999, p. 198-204.

Em resumo, a tecnologia é fruto da cultura.²³ Porém um velho sonho ainda impulsiona a imaginação humana rumo ao perigo reconfortante do desconhecido, o ser humano quer aprender mais, é uma curiosidade natural à espécie.

No campo da ciência da computação, Alan Turing²⁴, Von Neumann²⁵ e outros pesquisadores²⁶ foram os precursores do estudo acerca da Inteligência Artificial, indicando, de maneira visionária, um futuro que em parte já se vive. O tema da Inteligência Artificial passou por altos e baixos desde os grandes estudos publicados por Turing²⁷, muito devido às promessas não cumpridas de um futuro mirabolante²⁸, o que não foi uma exclusividade da Inteligência Artificial, haja vista que a tentativa de se antecipar o futuro era em meados do século XX (e ainda é nos dias atuais) um exercício marcado por imprecisões.²⁹

A cultura ocidental já vem se banhando nos pensamentos tecnológicos há algum tempo, como se observa nos robôs cheios de personalidade de Asimov em *Eu*,

²³ “No limite, o significado do termo tecnologia culmina em uma discussão sobre os aspectos sociais, econômicos e culturais de determinada sociedade em determinado contexto histórico.

Dessa forma, apesar de considerarmos hoje, em grande medida, tecnologia ‘sinônimo de aparelhos cada vez mais inteligentes, sofisticados e rápidos, como seu computador, tablet ou smartphone’, há quem diga que não é errado considerar ‘que um arco e flecha, por exemplo, também seja tecnologia’ (MAGRANI, Eduardo. **A Internet das Coisas**. Rio de Janeiro: FGV Editora, 2018, p. 44).

²⁴ TURING, Alan Mathison. Computing Machinery and Intelligence. **Mind**, n. 49, p. 433-460, 1950.

²⁵ VON NEUMANN, John. **The Computer and the Brain**. New Haven: Yale University Press, 1958.

²⁶ Cujos nomes principais no período entre 1943 e 1979 são: Warren McCulloch, Walter Pittes, Donald Webb, Marvin Minsky, Dean Edmonds, John McCarthy, Claude Shannon, Nathaniel Rochester, Trenchard More, Arthur Samuel, Ray Solomonoff, Oliver Selfridge, Carnegie Tech, Allen Newell, Herbert Simon, Daniel Brobow, David Huffman, David Waltz, Patrick Winston, Terry Winograd, Scott Fahlman, Bernie Widrow, Frank Rosenblatt, Ed Feigenbaum, Bruce Buchanan, Eugene Charniak e Roger Schank. (RUSSELL, Stuart; NORVIG, Peter. **Inteligência Artificial**. Tradução de Regina Célia Simille. Rio de Janeiro: Elsevier, 2013, p. 16-22).

²⁷ TURING, Alan M. On computable numbers, with an application to the *Entscheidungs* problem. **Journal of Math**, v. 58, 1936, p. 345-363.

²⁸ “O fraco desempenho desses primeiros sistemas também se devia a diversos outros fatores, tais como: métodos inadequados para lidar com incertezas, dependência de representações simbólicas frágeis e infundadas, escassez de dados e severas limitações de hardware tanto em capacidade de memória quanto em velocidade de processamento. Em meados da década de 1970, já havia uma conscientização crescente acerca desses problemas. A constatação de que muitos dos projetos de IA jamais poderiam realizar suas promessas iniciais levou ao surgimento do primeiro ‘inverno da IA’: um período de retração durante o qual os financiamentos diminuíram, o ceticismo aumentou, e a IA, então, saiu de moda” (BOSTROM, Nick. Op. cit., posição 456).

²⁹ “Uma queixa recorrente desses visionários impacientes é que hoje nos preocupamos com as questões erradas. Temos medo da ruptura, mas devíamos recebê-la de braços abertos. Em vez de evitar um salto no escuro, devíamos reconhecê-lo como a precondição de qualquer mudança significativa”. (...) “Poderíamos sugerir como lema político do século XXI: mais depressa!

Essa visão do mundo ganhou um rótulo filosófico próprio: aceleracionismo. Teve um precursor no início do século XX, o futurismo, que floresceu no período imediatamente anterior e posterior à Primeira Guerra Mundial. O futurismo celebrava a velocidade, a mecanização e a juventude” (RUNCIMAN, David. **Como a Democracia Chega ao Fim**. Tradução: Sergio Flaksman. São Paulo: Todavia, 2018, p. 212).

Robô (1950); no olhar desumano e impassível de *HAL 9000* em *2001: Uma Odisseia no Espaço* (1968); nos replicantes transgressores de *Blade Runner* (1982); na implacável *Skynet* de *O Exterminador do Futuro* (1984); no império de máquinas que escravizam corpo e mente humanas em *Matrix* (1999); em *David* a criança robô cheia de humanidade em *I.A. - Inteligência Artificial* (2001); no megalomaniaco *Ultron* e no fiel *J.A.R.V.I.S.* do universo da Marvel³⁰; na máquina que possui uma sedutora personalidade feminina em *Ex Machina* (2014); e na alma humana na androide de *Ghost In The Shell* (2017). Os exemplos são muitos e a lista referida está longe de ser exaustiva.

Nesse momento da história, antigos sonhos (e pesadelos) se materializam³¹, especialmente no que refere ao desenvolvimento da Inteligência Artificial. Ainda, um leque de ferramentas tecnológicas se abre e isso implica em efeitos diretos no modo como o direito é interpretado e aplicado, o que leva a efeitos reflexos na organização estatal. Assim, a tecnologia, aliada à jurisdição, vem evoluindo desde a bela caligrafia e a intermitente batida de teclas das máquinas de escrever³² e passou a dar destaque aos microchips.

A **primeira evolução** tecnológica da prestação jurisdicional passou pela utilização do **microcomputador** (que se avolumou no Brasil em meados da década de 1990). Se nos primórdios os juízes escreviam suas decisões à mão (e os servidores costuravam as folhas do processo), o que foi melhorado com a máquina de escrever, com o computador uma maior eficiência foi implementada.

³⁰ Os personagens *Ultron* e *J.A.R.V.I.S.* (“*Just A Rather Very Intelligent System*”) ficaram famosos com os filmes das franquias *Homem de Ferro* e *Vingadores*, mas suas origens naturalmente remontam respectivamente aos quadrinhos em 1968 (*Ultron*) e 2008 (quando a inteligência artificial foi batizada em homenagem ao clássico personagem Edwin Jarvis).

³¹ “Mas a revolução digital pode ter alterado os termos desse argumento. A internet das coisas desempenha um papel tão importante nas visões aceleracionistas do futuro porque promete libertar tanto as máquinas dos homens quanto os homens das máquinas. Permitirá que a mudança ocorra à velocidade das máquinas. Aos humanos, bastará se aproveitar do que as máquinas tornarem possível. Não precisamos mudar. Só precisamos embarcar na viagem da mudança.

Por isso, o perigo real pode ser irmos além, e não ficarmos aquém, do ponto para onde pensamos que estamos indo. A maioria das concepções utópicas de um mundo de redes liberadas supõe que os indivíduos que constituírem essas redes permanecerão como são hoje. Ainda poderemos escolher como e com quem queremos coexistir no panorama ilimitado que a internet descortinou para nós. Mas é possível que esse panorama se revele profundamente inóspito para a identidade humana individual” (RUNCIMAN, David. Op. cit., p. 213-214).

³² Qualquer leitor que tenha vivido os tempos anteriores à virada do milênio consegue ouvir, com alguma nostalgia, o tilintar de teclas, sinos e a engrenagens das máquinas escrever mecânicas, ou mesmo a robótica escrita das máquinas elétricas.

A **internet** (cuja ampliação de acesso ocorreu no final da década de 1990) aporta como a **segunda evolução** tecnológica do Judiciário³³, seu desenvolvimento e a melhoria da qualidade da rede (com velocidade de transmissão de dados) possibilitou a melhoria da comunicação dos tribunais com os jurisdicionados. Por exemplo, o acesso ampliado aos repositórios de julgados dos Tribunais possibilitou o aprimoramento do acesso à Justiça e maior transparência das decisões judiciais.

Já a **terceira evolução** se deve ao desenvolvimento dos **sistemas eletrônicos de auxílio à Justiça**³⁴ (Bacenjud³⁵, Renajud, Infojud, Serasajud, etc.), os quais representam o esforço na criação de soluções para a melhoria da atividade jurisdicional por meio do impacto da tecnologia, renovado modelo de organização das atividades jurisdicionais que diminuiu o martírio processual (das partes e do Juízo), que antes consistia nas rotinas de expedição de ofícios de papel para toda e qualquer consulta realizada pelo Poder Judiciário (evidencia-se que os sistemas de “ofícios eletrônicos” foram o primeiro grande passo tecnológico do Poder Judiciário).

A **quarta evolução** advém do aprimoramento tecnológico inerente ao **processo eletrônico** (o qual foi regulamentado no Brasil pela Lei 11.419/2006), sendo que, desde 1997 (e principalmente durante a década de 2000) os Tribunais que no início apresentavam resistência à mudança de paradigma tecnológico foram paulatinamente desenvolvendo seus próprios sistemas processuais (muitas vezes antes de existir legislação processual sobre o tema) ou adotando modelos já prontos. Merecem destaque o pioneiro **e-SAJ** (adotado pelos Tribunais de Justiça de Santa Catarina – em 1997³⁶ – e São Paulo³⁷, bem como por outros Tribunais), o **Eproc**³⁸ (da

³³ MAGRANI, Eduardo. **A Internet das Coisas**, p. 63-70.

³⁴ CONSELHO NACIONAL DE JUSTIÇA. Recomendação nº 51, de 24 de março de 2015. Recomenda a utilização dos Sistemas Bacenjud, Renajud e Infojud e dá outras providências. **Diário da Justiça**: Brasília, DF, n.º 54/2015, p. 3, 24 mar. 2015.

³⁵ CONSELHO NACIONAL DE JUSTIÇA. Resolução nº 61, de 07 de outubro de 2008. Disciplina o procedimento de cadastramento de conta única para efeito de constrição de valores em dinheiro por intermédio do Convênio BACENJUD e dá outras providências. **Diário da Justiça**: Brasília, DF, n.º 68/2008, p. 2-4, 15 out. 2008.

³⁶ SANTA CATARINA. Tribunal de Justiça de Santa Catarina. **Processo Eletrônico SAJ: Histórico**. [2020]. Disponível em: <https://www.tjsc.jus.br/web/processo-eletronico-saj/historico>. Acesso em 01 mai. 2020.

³⁷ SÃO PAULO. Tribunal de Justiça do Estado do Paraná. Resolução nº 551, de 31 de agosto de 2011. Regulamenta o processo eletrônico no âmbito do Tribunal de Justiça do Estado de São Paulo e dá outras providências. **Diário da Justiça Eletrônico**: São Paulo, SP, Ano IV, edição n.º 1033, 08 set. 2011.

³⁸ BRASIL. Tribunal Regional Federal da 4ª Região. **Eproc: processo eletrônico da Justiça Federal da 4ª Região completa cinco anos**. [2014]. Disponível em: https://www.trf4.jus.br/trf4/controlador.php?acao=noticia_visualizar&id_noticia=10500. Acesso em 01 mai. 2020.

Justiça Federal³⁹), o **Projudi**⁴⁰ (atualmente adotado por dezenove Tribunais de Justiça⁴¹), o **PJe**⁴² (da Justiça do Trabalho⁴³, embora inicialmente desenvolvido pelos Tribunais Regionais Federais em parceria com o Conselho Nacional de Justiça⁴⁴, hoje também utilizado no âmbito do próprio CNJ), o **E-STJ**⁴⁵ e o **e-STF**⁴⁶.

Por fim, a **quinta evolução** tecnológica dos serviços jurisdicionais está ligada ao desenvolvimento e à aplicação de **Inteligência Artificial**⁴⁷ computacional às variadas atividades inerentes ao Poder Judiciário (desde uma simples contagem de prazo, passando por atos de comunicação processual e, até mesmo, na decisão judicial⁴⁸), o que traz consigo colossais promessas de melhoria⁴⁹ (celeridade,

³⁹ A Resolução nº 13/2004, do Tribunal Regional Federal da 4ª Região determinou o início do processo eletrônico no âmbito dos Juizados Especiais Federais, sendo que após verificada a eficiência do sistema houve a extensão para toda a Justiça Federal a 4ª Região, por meio da Resolução nº 64/2009, do TRF4. (BRASIL. Tribunal Regional Federal da 4ª Região. Resolução nº 13, de 11 de março de 2004. Implanta e estabelece normas para o funcionamento do Processo Eletrônico nos Juizados Especiais Federais no âmbito da Justiça Federal da 4ª Região. **Boletim Interno Eletrônico**: Porto Alegre, RS, n.º 174-06, 20 abr. 2004; BRASIL. Tribunal Regional Federal da 4ª Região. **Resolução nº 64, de 11 de novembro de 2009**. Implanta o Processo Eletrônico no âmbito da Justiça Federal de 1º e 2º Graus da 4ª Região. Porto Alegre. TRF4, 11 nov. 2009. Disponível em: <https://www.trf4.jus.br/trf4/sup/res64-2009.pdf>. Acesso em: 02 fev. 2022).

⁴⁰ Por exemplo: PARANÁ. Tribunal de Justiça do Estado do Paraná. Resolução nº 03, de 14 de maio de 2009. Dispõe sobre o processo eletrônico no âmbito do Poder Judiciário do Estado do Paraná. **Diário da Justiça Eletrônico**: Curitiba, PR, n.º 138, 14 mai. 2009.

⁴¹ CONSELHO NACIONAL DE JUSTIÇA. **Área de Atuação do Projudi**. [2020]. Disponível em: <https://www.cnj.jus.br/area-de-atuacao-projudi>. Acesso em: 01 mai. 2020.

⁴² BRASIL. Tribunal Superior do Trabalho. **Histórico do PJe**. [2020]. Disponível em: <http://www.tst.jus.br/web/pje/historico>. Acesso em 01 mai. 2020.

⁴³ CONSELHO NACIONAL DE JUSTIÇA; CONSELHO SUPERIOR DA JUSTIÇA DO TRABALHO. **Acordo de Cooperação Técnica nº 01/2010**. Acordo de Cooperação Técnica que entre si celebram o Tribunal Superior do Trabalho, o Conselho Superior da Justiça do Trabalho e os Tribunais Regionais do Trabalho. Brasília: CNJ e CSJT, 29 mar. 2010. Disponível em: <https://www.csjt.jus.br/documents/955023/9626845/Desenvolvimento+do+sistema+de+processo+judicial+e+letr%C3%B4nico.pdf/d54be507-cabc-73c5-ab7c-2dd8d07a8c97?t=1635361011200>. Acesso em: 02 fev. 2022.

⁴⁴ CONSELHO DA JUSTIÇA FEDERAL. **Histórico do PJe**. [2020]. Disponível em: https://www.cjf.jus.br/observatorio/arq/historico_pje.pdf. Acesso em: 01 mai. 2020.

⁴⁵ BRASIL. Superior Tribunal de Justiça. Resolução STJ/GP nº 10, de 6 de outubro de 2015. Regulamenta o processo judicial eletrônico no Superior Tribunal de Justiça. **Diário de Justiça Eletrônico do STJ**: Brasília, DF, 07 out. 2015.

⁴⁶ BRASIL. Supremo Tribunal Federal. Resolução nº 427, de 20 de abril de 2010. Regulamenta o processo eletrônico no âmbito do Supremo Tribunal Federal. **Diário de Justiça Eletrônico do STF**: Brasília, DF, nº 72, p. 1, 26 abr. 2010.

⁴⁷ HARTMANN PEIXOTO, Fabiano; SILVA, Roberta Zumblick Martins da. **Inteligência Artificial e Direito**. Vol. 1. Curitiba: Alteridade, 2019, p. 119-121.

⁴⁸ SARTOR, Giovanni. Doing justice to rights and values: teleological reasoning and proportionality. **Artificial Intelligence and Law**, Dordrecht, v. 18, p. 175-215, 2010. DOI: 10.1007/s10506-010-9095-7.

⁴⁹ *"There is a profound message here for lawyers—when thinking about technology and the internet, the challenge is not just to automate current working practices that are not efficient. The challenge is to innovate, to practise law in ways that we could not have done in the past"* (SUSSKIND, Richard. **Tomorrow's Lawyers**: An Introduction to Your Future. 2ª ed. Oxford: Oxford University Press, 2017, p. 14-15. E-book Kindle).

objetividade, menor custo, maior eficiência, etc.), de um lado, e severas críticas (vieses decisórios, ausência de transparência, risco de desumanização das decisões, possível congelamento jurisprudência, etc.), por outro.

Nesse sentido, o presente estudo se pauta numa visão equilibrada da IA⁵⁰, a fim de que haja, com serenidade e seriedade, o desenvolvimento de mecanismos de aprimoramento da função jurisdicional, dentro de um pensar de que a inafastabilidade da jurisdição (art. 5º, XXXV, da Constituição Federal) deve se reinventar tecnologicamente, sob pena de ser suprimida por outras soluções⁵¹ não estatais de tutela de direitos⁵², pois o Estado Democrático é um modelo que vem sendo desafiado por modelos concorrentes⁵³, o que não é diferente no que toca ao Judiciário.

Ademais, percebe-se que, para que haja a evolução continua e exponencial⁵⁴ do processo eletrônico e da Inteligência Artificial nele a ser integrada (*embedded*⁵⁵), uma aresta institucional que se formou nos últimos anos deve ser aparada, a relação entre os Tribunais e o Conselho Nacional de Justiça⁵⁶ no aspecto de gestão de tecnologia. Apesar de todas as tentativas de se adotar o PJ-e como modelo nacional de processo eletrônico (Resolução nº 185/2013 do CNJ), muitos Tribunais optaram por desenvolver modelo próprio ou por adotar outros modelos para os seus processos eletrônicos (com destaque para o TJSC, que na prática foi impedido pelo CNJ de adotar o sistema Eproc⁵⁷). A esse respeito a regra do art. 8º da Lei 11.419/2006 prevê que os “órgãos do Poder Judiciário **poderão desenvolver sistemas eletrônicos de processamento de ações judiciais** por meio de autos total ou parcialmente digitais...” [destaque nosso], limitando, no caso, o CNJ a regular os padrões de qualidade e eficiência dos sistemas de processo eletrônico desenvolvidos

⁵⁰ HARTMANN PEIXOTO, Fabiano. **Inteligência Artificial e Direito**: Convergência Ética e Estratégica. Vol. 5. Curitiba: Alteridade, 2020, p. 26.

⁵¹ BOSTROM, Nick. Op. cit., posição 2779-2784.

⁵² RUNCIMAN, David. Op. cit., p. 143-146.

⁵³ Este debate, apesar de bastante interessante, é maior que os limites propostos para a presente pesquisa.

⁵⁴ ISMAIL, Salim; VAN GEES, Yuri; MALONE, Michael S. **Organizações exponenciais**. Rio de Janeiro: Alta Books, 2018, posição 331-350. E-book Kindle.

⁵⁵ GREENFIELD, Adam. **Everyware**: The dawning age of ubiquitous computing. Berkeley: Pearson Education, 2006, p. 134. E-book Kindle.

⁵⁶ O CNJ tem importantes atribuições constitucionais (art. 103-B, §4º, da CF), especialmente no que toca ao necessário aperfeiçoamento administrativo (em temas como gestão, carreiras de servidores e responsabilidade disciplinar, por exemplo), o que não lhe deu a possibilidade de frear o desenvolvimento tecnológico do Judiciário em suas diversas frentes.

⁵⁷ CONSELHO NACIONAL DE JUSTIÇA. **TJSC firma acordo e implantará sistema eletrônico do CNJ**. [2020]. Disponível em: <https://www.cnj.jus.br/tjsc-firma-acordo-e-implantara-sistema-eletronico-do-cnj>. Acesso: 01 mai. 2020.

(art. 103-B, §4º, I e II, da CF e art. 196 do CPC), sem que isso implique na obstrução da aplicação da referida regra da Lei de Processo Eletrônico⁵⁸, uma vez que a palavra “poderão” não está inserida na norma legal por acidente. Em suma, como não se cogita de inconstitucionalidade no art. 8º da Lei 11.419/2006, este deve ser seguido fielmente pelo CNJ e pelos Tribunais, sob pena de soluções locais inovadoras na matéria de Inteligência Artificial jurídica serem desmotivadas por modelo de extrema centralização.

A política de centralização do desenvolvimento de sistemas de processo eletrônico não pode ser repetida no que tange aos sistemas de Inteligência Artificial processual. O processo eletrônico, pautado em Inteligência Artificial e no paradigma de multiplataforma (com “independência da plataforma computacional” e “interoperabilidade dos sistemas” – art. 194 do CPC), ainda tem muitíssimos aspectos a serem desenvolvidos⁵⁹, uma vez que, para além da ilegalidade, percebe-se a ausência de razoabilidade em qualquer tentativa de obstrução de que cada Tribunal tenha sistema próprio de processo baseado em Inteligência Artificial. Nesse sentido, o que deve ser garantido no plano regulatório nacional é a interoperabilidade entre os diversos sistemas e a existência de um cadastro único de perfis de acesso para todos os Tribunais brasileiros, o que facilita o acesso à Justiça por advogados (e partes) de qualquer parte do país, mediante efetiva desburocratização.

Para além disso, destaca-se que uma solução geral pronta e “enlatada” para o desenvolvimento de um processo eletrônico totalmente robotizado pode ser perigosa, haja vista que numa perspectiva democrática as aplicações baseadas em IA devem ter, em suas fases de desenvolvimento, a previsão de margem de adaptabilidade a cada unidade jurisdicional e à arquitetura de decisões pensada pelo magistrado que supervisiona a unidade. Na prática, podem surgir críticas no sentido de que a proposta é de difícil implementação, porém esse é um desafio que o Direito deve fazer aos profissionais da tecnologia, pois a IA judicial nasce com limitações legais e constitucionais intrínsecas ao adjetivo. A IA aplicada à jurisdição se converte num

⁵⁸ Esta regra foi legitimamente advinda de democrática deliberação do Poder Legislativo no uso de sua competência legislativa privativa prevista no art. 22, I, da Constituição Federal.

⁵⁹ ALVES, Isabella Fonseca; ALMEIDA, Priscila Brandão de. Direito 4.0: uma análise sobre inteligência artificial processo e tendências de mercado. *In*: ALVES, Isabella Fonseca (org.). **Inteligência Artificial e Processo**. Belo Horizonte: D’Plácido, 2020, p. 57-68.

instrumento pautado na racionalidade⁶⁰ inerente à função democrática⁶¹, não se voltando simplesmente à substituição dessa.

Dessa forma, embora iniciativas como a Resolução nº 332/2020 do Conselho Nacional de Justiça⁶² tenham estimulado o modelo colaborativo, referido órgão constitucional-administrativo não estabeleceu o paradigma da IA centrada no modelo de independência judicial.⁶³

A IA judicial deve ser desenvolvida mediante *design by judge*, uma leitura possível a partir do *design* dirigido ao domínio (“*domain-driven design*”⁶⁴).

A tendência de desenvolvimento colaborativo⁶⁵, entre os Tribunais, o terceiro setor e a iniciativa privada, de aplicações suplementares que funcionem em plena interface com os sistemas de processo eletrônico deve ser política pública nacional do Poder Judiciário⁶⁶, a fim de que se evite o estrangulamento da inventividade de novas soluções em IA a serem propostas e, ao mesmo tempo, haja um campo seguro para a manutenção de sistemas já operacionais e em vias de implementação pelos Tribunais. Inovar é um mandamento que deverá nortear o Judiciário, devendo existir a consciência que a computabilidade do direito é desafio para ser estudado em várias

⁶⁰ “Quem se engana sistematicamente sobre si mesmo comporta-se de maneira irracional. No entanto, quem está em condições de aceitar esclarecimentos sobre a sua irracionalidade dispõe da racionalidade de um sujeito capaz de emitir juízos e agir de modo que se oriente racionalmente a um fim; dispõe de racionalidade de um sujeito capaz de discernir moralmente, que se revela confiável do ponto de vista prático, que atribui valor com sensibilidade e está esteticamente aberto; e mais que isso: alguém assim dispõe sobretudo da energia necessária para se comportar de maneira reflexiva diante da própria subjetividade e para entrever as limitações irracionais a que estão sistematicamente submetidas suas exteriorizações cognitivas, prático-morais e prático-estéticas” (HABERMAS, Jürgen. **Teoria do Agir Comunicativo, 1: Racionalidade da Ação e Racionalidade Social**. Tradução: Paulo Astor Soethe. São Paulo: WMF Martins Fontes, 2012, p. 54).

⁶¹ “O princípio do discurso tem inicialmente o *sentido cognitivo* de filtrar contribuições e temas, argumentos e informações, de tal modo que os resultados obtidos por este caminho têm a seu favor a suposição de aceitabilidade racional: o procedimento democrático deve fundamentar a legitimidade do direito”. (...) “E esse cruzamento entre *normatização discursiva do direito e formação comunicativa do poder* é possível, em última instância, porque no agir comunicativo os argumentos também formam motivos” (HABERMAS, Jürgen. **Direito e Democracia: entre facticidade e validade**. Vol. I. Tradução de Flávio Beno Siebeneichler. Brasília: Tempo Brasileiro, 1997, p. 191).

⁶² CONSELHO NACIONAL DE JUSTIÇA. Resolução nº 332, de 21 de agosto de 2020. Dispõe sobre a ética, a transparência e a governança na produção e no uso de Inteligência Artificial no Poder Judiciário e dá outras providências. **Diário da Justiça**: Brasília, DF, n.º 274, p. 4-8, 25 ago. 2020.

⁶³ A esse respeito a seção 63 da *Recomendação sobre Ética da Inteligência Artificial*, da UNESCO, prevê o necessário respeito à independência judicial nos sistemas de IA a serem utilizados no Poder Judiciário, bem como o princípio da supervisão humana para garantir um desenvolvimento orientado e centrado no ser humano e no uso de sistemas de IA no judiciário (UNESCO. **Recommendation on the Ethics of Artificial Intelligence**. 24 Nov 2021. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000380455>. Acesso em: 20 mar. 2022).

⁶⁴ EVANS, Eric. **Domain-Driven Design: Atacando as Complexidades no Coração do Software**. 3ª ed. Tradução: Tibério Júlio Couto Novais. Rio de Janeiro: Alta Books, 2016, p. 198.

⁶⁵ O CNJ possui modelo colaborativo vinculado ao sistema PJe.

⁶⁶ E este é um ponto de relativo acerto da Res. 332/2020 do CNJ.

frentes, sendo inviável o monopólio do conhecimento ou o aprisionamento da tecnologia por quem quer que seja.

Assim, o serviço jurisdicional atingirá maior grau de respeito ao princípio constitucional da eficiência (art. 37, *caput*, do Constituição Federal⁶⁷), o que sempre se esperou do processo eletrônico.

1.2 Inteligência Artificial: Aspectos de Um Conceito Não Jurídico

Há uma tentação dos juristas em definir tudo através de conceitos, inclusive o que é (ou não é) Inteligência Artificial. Contudo, o conceito Inteligência Artificial é disputado na própria Ciência da Computação. Existe quem entenda que o termo IA deva ser reservado apenas a programas de computador dotados de grande autonomia (ainda que não possam ser considerados uma IA de tipo geral). Existem outras posições que defendem que a execução de tarefas de forma automatizada, mediante um sistema altamente especializado⁶⁸, já possibilitaria a denominação do fenômeno como baseado em Inteligência Artificial.⁶⁹

Segundo Stuart Russell e Peter Norvig, o estudo da Inteligência Artificial passa por quatro correntes de pensamento na Ciência da Computação, as quais consistem em desenvolver sistemas computacionais voltados a qualquer dos seguintes aspectos: a) **pensar como um humano**; b) **agir como um humano**; c) **pensar**

⁶⁷ O qual é representado no âmbito especificamente processual pelo princípio da eficiência previsto no art. 8º do Código de Processo Civil.

⁶⁸ “Anteriormente se supunha, não totalmente sem razão, que para um computador jogar xadrez no nível de um grande mestre, teria que ser dotado com um alto grau de inteligência geral. Imaginava-se, por exemplo, que jogar xadrez requeria capacidade de aprender conceitos abstratos, pensar de forma inteligente sobre estratégia, compor planos flexíveis, empregar uma ampla gama de deduções lógicas complexas e talvez até mesmo modelar o pensamento do seu oponente. Mas isso não é verdade. Confirmou-se que é possível construir um mecanismo de xadrez perfeitamente bom em torno de um algoritmo de propósito específico. Quando esse algoritmo foi implementado nos processadores rápidos que se tornaram disponíveis no fim do século XX, resultou num excelente jogo. Mas uma IA construída dessa forma é limitada. Ela não faz nada além de jogar xadrez” (BOSTROM, Nick. Op. cit., posição 714).

⁶⁹ “No passado, havia muitas coisas que somente os humanos podiam fazer. Mas hoje robôs e computadores estão assumindo esse papel e logo poderão sobrepujar os humanos no cumprimento da maioria das tarefas. É verdade que o funcionamento dos computadores é muito diferente do dos humanos, e parece improvável que eles se tornem humanoides em pouco tempo. Em particular, não parece que computadores estejam prestes a ter consciência, nem emoções e sensações. As últimas décadas assistiram a um avanço imenso na inteligência dos computadores, mas o avanço na consciência dessas máquinas foi nulo. Até onde sabemos, computadores não são, em 2016, mais conscientes do que seus protótipos na década de 1950. No entanto, estamos à beira de uma grave revolução. Humanos correm o perigo de perder seu valor porque a inteligência está se desacoplando da consciência” (HARARI, Yuval Noah. **Homo Deus**: Uma breve história do amanhã. Tradução: Paulo Geiger. São Paulo: Companhia das Letras, 2016, p. 313).

racionalmente; ou d) **agir racionalmente**.⁷⁰ Desse modo, as técnicas de aprendizado de máquina (*machine learning*) e de redes neurais de aprendizado profundo (*deep learning*) fazem parte do estado da arte da Inteligência Artificial⁷¹, mas não deslegitimam o uso de conceito Inteligência Artificial para outras técnicas de IA.

Na concepção de Fabiano Hartmann Peixoto e Roberta Zumblick Martins da Silva, a Inteligência Artificial é um “termo guarda-chuva” que “abriga muitas áreas de estudos e técnicas, como visão computacional, robótica, processamento de linguagem natural e *machine learning*, por exemplo”.⁷² Nesse sentido, não se pode confundir IA apenas com algoritmos, “porque os sistemas de IA são compostos de algoritmos (livremente definidos como programas), juntamente com outras aplicações de computadores”.⁷³

Ademais, conforme será visto no Capítulo 5, especificamente em relação à IA baseada em aprendizado de máquina (*machine learning*), esse gênero de Inteligência Artificial é desenvolvido com a tríade **algoritmos, bases de dados (*datasets*) e modelos de *machine learning***.

Para além do debate sobre as características principais para um sistema computacional ser classificado como baseado em Inteligência Artificial (debate que é necessário do ponto de vista da ciência da computação e da filosofia), a presente pesquisa se pauta na ideia de que **a Inteligência Artificial depende da existência de grau de autonomia decisória inerente aos sistemas computacionais inteligentes**.⁷⁴ Grau de autonomia esse que pode ser maior ou menor, uma vez que é a mensuração do grau de autonomia da Inteligência Artificial que traz soluções e problemas para o mundo jurídico quando se pensa na aplicabilidade de IA ao Direito. Desse modo, o conceito de “modelo de inteligência artificial” presente no art. 3º, II, da Res. 332/2020 do CNJ⁷⁵ é mais estreito que o conceito tradicional de IA adotado na

⁷⁰ RUSSELL, Stuart; NORVIG, Peter. Op. cit., p. 3-6.

⁷¹ Ibid., p. 26-27.

⁷² HARTMANN PEIXOTO, Fabiano; SILVA, Roberta Zumblick Martins da. **Inteligência Artificial e Direito**, p. 75.

⁷³ Ibid., p. 76.

⁷⁴ DEWEY, Daniel; RUSSEL, Stuart; TEGMARK, Max. Research Priorities for Robust and Beneficial Artificial Intelligence. **AI Magazine**, v. 36, n. 4, 2015.

⁷⁵ “Modelo de Inteligência Artificial: conjunto de dados e algoritmos computacionais, concebidos a partir de modelos matemáticos, cujo objetivo é oferecer resultados inteligentes, associados ou comparáveis a determinados aspectos do pensamento, do saber ou da atividade humana” (CONSELHO NACIONAL DE JUSTIÇA. Resolução nº 332, de 21 de agosto de 2020).

ciência da computação, característica inerente à regulação do CNJ, mas que não é limitante à presente pesquisa ao se adotar o conceito amplo de IA.

Logo, o problema da existência de uma Inteligência Artificial do tipo geral, capaz de responder a uma espécie de *teste de Turing*⁷⁶ e se passar por um operador do direito (se passar por jurista seria mais difícil), não é preocupação para os limites da presente pesquisa, uma vez que os sistemas de IA jurídica até aqui desenvolvidos são altamente especializados em tarefas pontuais (organização de documentos, geração automatizada de documentos e na orientação jurídica on-line em campos específicos do conhecimento jurídico)⁷⁷ ou sem foco específico na interpretação jurídica de qualquer situação que possa surgir. Por exemplo, o pioneiro, e já descontinuado⁷⁸, sistema *ROSS Intelligence*⁷⁹ possuía limitações, pois era um ótimo assistente ao combinar processamento de linguagem natural do sistema *Watson* da IBM, *machine learning* e *e-discovery*, mas não se tratava uma máquina de interpretação do direito.

Dessa maneira, a pesquisa também leva em consideração a futura possibilidade de desenvolvimento de sistemas com Inteligência Artificial, especialmente através de *machine learning*⁸⁰ e redes neurais⁸¹ de aprendizado profundo (*deep learning*), e sua aplicabilidade à interpretação do direito, de maneira a dar um passo à frente em relação aos modelos anteriormente propostos (desde a automatização mediante programação procedural até a lógica bayesiana⁸²).

A UNESCO, ciente das dificuldades e disputas acerca do conceito de Inteligência Artificial, assim se posicionou sobre o tema em sua *Recomendação sobre a Ética da Inteligência Artificial*:

⁷⁶ TURING, Alan Mathison. *Computing Machinery and Intelligence*.

⁷⁷ SUSSKIND, Richard. **The End of Lawyers?** Rethinking the Nature of Legal Services. Oxford: Oxford University Press, 2010, p. 99-145. E-book Kindle.

⁷⁸ O *ROSS Intelligence* foi encerrado devido a conflitos jurídicos com a Thomson Reuters (*ROSS Intelligence. Announcement*. [2021]. Disponível em: blog.rossintelligence.com/post/announcement. Acesso em: 04 nov. 2021).

⁷⁹ *ROSS Intelligence. Ross Intelligence*. [2019]. Disponível em: <https://rossintelligence.com>. Acesso em: 31 ago. 2019.

⁸⁰ FERNANDES, W. P. D., SILVA, L. J. S., FRAJHOF, I. Z. *et al.* Appellate Court Modifications Extraction for Portuguese. **Artificial Intelligence and Law**, 2019. DOI: 10.1007/s10506-019-09256-x.

⁸¹ NGUYEN, T., NGUYEN, L., TOJO, S. *et al.* Recurrent neural network-based models for recognizing requisite and effectuation parts in legal texts. **Artificial Intelligence and Law**, v. 26, p. 169–199, 2018. DOI: 10.1007/s10506-018-9225-1.

⁸² GILL, Richard David; FENTON, Norman; LAGNADO, David; NEIL, Martin. Modelling Competing Legal Arguments Using Bayesian Model Comparison and Averaging. **Artificial Intelligence and Law**, v. 27, p. 403-430, 2019. DOI: 10.1007/s10506-019-09250-3.

Esta Recomendação não tem a ambição de fornecer uma definição única de IA, uma vez que tal definição precisaria mudar ao longo do tempo, de acordo com a evolução tecnológica. Em vez disso, sua ambição é abordar os recursos dos sistemas de IA que são de relevância ética central.⁸³

Em resumo, guardada a ideia de **grau de autonomia decisória** típico dos modelos de IA adotada neste estudo, não é tarefa do Direito definir o conceito de Inteligência Artificial.

1.3 Alvorecer: Inteligência Artificial e Direito

No Direito, pode-se citar que desde 1987, a *International Association for Artificial Intelligence and Law*⁸⁴ vem realizando uma compilação das pesquisas acerca da relação entre o direito e a Inteligência Artificial, sendo que, naturalmente, as pesquisas avançam de maneira a acompanhar as novas tecnologias.

Logo, a preocupação da integração da tecnologia à aplicação do direito não é algo tão recente quanto possa aparentar. Fora do Brasil, Kevin D. Ashley⁸⁵ e diversos outros pesquisadores⁸⁶ vêm trabalhando com afinco no tema desde o final dos anos 1980. No espaço nacional, pode-se citar o trabalho de pesquisadores como Aires José Rover⁸⁷ e José Renato Gaziero Cella.⁸⁸ Contudo, nos últimos anos, as soluções de

⁸³ Tradução livre de: “*This Recommendation does not have the ambition to provide one single definition of AI, since such a definition would need to change over time, in accordance with technological developments. Rather, its ambition is to address those features of AI systems that are of central ethical relevance*” (UNESCO. Op. cit. p. 3).

⁸⁴ Ver mais em www.iaail.org.

⁸⁵ ASHLEY, Kevin D. Case-based reasoning and its implications for legal expert systems. **Artificial Intelligence and Law**, Dordrecht, v. 1, p. 113-208, 1992, DOI 10.1007/BF00114920.

⁸⁶ BENCH-CAPON, Trevor et al. A history of AI and Law in 50 papers: 25 years of the international conference on AI and Law. **Artificial Intelligence and Law**, Dordrecht, v. 20, p. 215-319, 2012. DOI: 10.1007/s10506-012-9131-x.

⁸⁷ ROVER, Aires José. Sistemas Especialistas Legais: Pensando o Sistema Jurídico. **Sequência, estudos jurídicos e políticos**, Florianópolis, v. 29, p. 65-79, dez. 1994.

⁸⁸ CELLA, José Renato Gaziero. **Controle das Decisões Jurídicas pela Técnica do Autoprecedente: Lógica Deôntica Paraconsistente Aplicada em Sistemas Especialistas Legais**. Tese (Doutorado em Direito) – Universidade Federal de Santa Catarina, Florianópolis, p. 268, 2008.

Inteligência Artificial baseadas em modelos de *machine learning*⁸⁹ e *deep learning*⁹⁰ alcançaram níveis antes não experimentados, o que trouxe a possibilidade de maior aplicabilidade ao direito, a qual vem acompanhada de grandes preocupações.

Em meio ao ambiente tecnológico pode o jurista se embrenhar nesse campo do estudo multidisciplinar que envolve direito, ciência da computação, engenharia da computação, lógica, matemática e linguística. Isso é um desafio que, por si só, marcará a experiência do jurista que tiver a coragem de fazê-lo. A meta é olhar para o abismo, ainda que ele olhe de volta para você. De plano, é necessária uma abordagem que supere três espécies de posição que podem ser consideradas reticentes à simbiose⁹¹ entre Inteligência Artificial e teoria da decisão judicial.

A primeira posição pode ser denominada como o **desalento com o novo**, que parte da premissa que a Inteligência Artificial implicaria num afastamento dos operadores do direito das atividades inerentes à aplicação e à interpretação jurídicas, destruindo a essência daquilo que o direito sempre foi. Uma ideia típica dessa classe de pensamento é a de que o direito agora será dilacerado pela tecnologia, que a figura do juiz perderá a sua função e que a pesquisa em Inteligência Artificial seria, ao final, tempo perdido aos juristas.

A segunda posição pode ser denominada **alarmismo apocalíptico**, ela se pauta numa visão de que toda Inteligência Artificial poderia se voltar contra o homem, vindo a substituir a humanidade como espécie dominante ou a escravizá-la.⁹² A inspiração,

⁸⁹ “A kind of ML that has been applied to predict legal outcomes is supervised ML. Since it involves inferring a classification model (or function) from labeled training data, the ML is referred to as supervised (see Mohri et al., 2012).

The training data comprise a set of examples that have been assigned outcomes. Each example is a pair consisting of an input object (often a vector of feature values) and a desired output value. The learning algorithm needs to generalize from the training data to unseen situations” (ASHLEY, Kevin D. **Artificial Intelligence and Legal Analytics**: New tools for law practice in digital age. Cambridge: Cambridge University Press, 2017, p. 109).

⁹⁰ “*Deep learning* é uma forma específica de *machine learning*, que envolve o treinamento de redes neurais com muitas camadas de unidades. Ficou popular nos últimos anos e levou a melhorias significativas em tarefas de reconhecimento visual de objetos e reconhecimento de discurso” (HARTMANN PEIXOTO, Fabiano; SILVA, Roberta Zumblick Martins da. Op. cit., p. 99).

⁹¹ O termo “simbiose” pressupõe que os operadores do direito de carne e osso ainda continuem necessários.

⁹² “Entretanto, havia dois grandes riscos em transformar o Estado num autômato gigante. O primeiro: ele não ser suficientemente poderoso. Outras criaturas artificiais que fossem mais implacáveis, mais eficientes, mais robóticas — e, por implicação, menos humanas — poderiam resultar mais fortes. O segundo: que ele acabasse parecido demais com as coisas que fora criado para regular. Num mundo de máquinas, o Estado poderia adotar o caráter dominante e se tornar totalmente artificial. Este é o medo original da era moderna: não o que acontece quando as máquinas ficam parecidas demais conosco, mas quando ficamos parecidos demais com as máquinas” (RUNCIMAN, David. Op. cit., p. 138).

nesse caso, é bastante marcada pela influência de obras literárias e cinematográficas de ficção e por um neoludismo⁹³ pautado no receio de que todas as atividades humanas sejam substituídas por Inteligência Artificial.

A terceira posição pode ser chamada de **arrebato tecnológico**, a qual se opõe às primeiras posições na medida em que defende a ideia de que a tecnologia (no que se enquadra a Inteligência Artificial) tem um viés de salvamento da humanidade de suas agruras mundanas, haja vista que assim o homem poderia ter mais tempo para se dedicar às atividades intelectuais, teria que trabalhar menos e conseguiria aproveitar melhor a vida (inclusive através do aprimoramento biotecnológico). Em última instância, o homem passaria a subsistir num artificioso estado de paz tecnológica.⁹⁴

Considerando cada uma das posições aqui resumidas, é evidente que as mesmas trazem pontos de vista interessantes e parcialmente válidos, devendo-se buscar um delicado equilíbrio entre as três posições referidas, evitando cair nas armadilhas que cada linha de pensamento traz, bem como buscando compreender as visões preditivas do futuro. Neste momento, de incertezas tecnológicas aos estudiosos do Direito, cabe tatear o passado com a humildade que só a experiência traz, inspecionar o presente com os olhos atentos à realidade mutante e sondar as frestas do tempo para buscar aquilo que o futuro reserva.

Desse modo, ainda que a IA já venha desempenhando papel de destaque há algumas décadas, especialmente os estudos e a aplicação da Inteligência Artificial vêm se avolumando e se acelerando, o que pode ser sentido nos estudos atuais acerca da aplicação de *machine learning*, *deep learning* e processamento de

⁹³ KURZWEIL, Ray. **A Singularidade Está Próxima**: Quando os Humanos Transcendem a Biologia. Tradução: tradução Ana Goldberger. São Paulo: Iluminuras, 2018, p. 473.

⁹⁴ “ *The pax technica is a political, economic, and cultural arrangement of social institutions and networked devices in which government and industry are tightly bound in mutual defense pacts, design collaborations, standards setting, and data mining. Over the past quarter-century we have learned a lot about the political impact of new information technologies. Given that our internet is evolving into an internet of connected things, it makes sense to apply what we know in some conservative premises. If these are the premises of the pax technica, what are the consequences—desirable or otherwise?*

We have begun an extended period of stability brought about by the dominance of an internet built and maintained by Western democracies. This expansive network of devices has allowed viral social movements that are massive, networked, leaderless, temporary, and multi-issue. Power flows to people and organizations that control digital media or that do creative things with information infrastructure” (HOWARD, Philip N. **Pax Technica**: How the Internet of Things May Set Us Free or Lock Us Up. New Haven: Yale University Press, 2015, p. 145-146. E-book Kindle).

linguagem natural para soluções em voltadas à produção de documentos⁹⁵ (contratos, petições e minutas de decisões judiciais).

Existem diversos aspectos e aplicabilidades da IA junto à atividade jurisdicional, como o caso das atividades de movimentação processual, controle de prevenção de nulidades processuais, conciliação, mediação, classificação de processos e elaboração de minutas de decisões (judiciais ou de arbitragem).

Nesse ponto, a aplicação do direito nas decisões judiciais pode se beneficiar de diversas tecnologias de Inteligência Artificial (já desenvolvidas ou a serem implementadas⁹⁶), uma vez que essa é a fronteira a ser perseguida na relação entre os campos eminentemente jurídicos (teoria da interpretação e teoria da decisão) e a Inteligência Artificial (a qual é dotada de subcampos próprios de estudo).

Destaca-se que, no Brasil, a função jurisdicional é exclusiva dos juízes, integrantes de carne e osso da carreira da magistratura, na forma definida pelo art. 93 da Constituição Federal de 1988, os quais têm o dever funcional de independência em relação a qualquer tipo de influência externa às suas decisões (art. 35, I, da Lei Complementar 35/1979⁹⁷ e do art. 4º ao art. 7º do Código de Ética da Magistratura⁹⁸). Nesse aspecto, o art. 5.º do Código de Ética da Magistratura é bastante claro ao prever que deve o “magistrado pautar-se no desempenho de suas atividades sem receber indevidas influências externas e estranhas à justa convicção que deve formar para a

⁹⁵ LINNA JR., Daniel W. What We Know and Need to Know About Legal Startups. **South Carolina Law Review**, V. 67, p. 389-417, 2016.

⁹⁶ “A Inteligência Artificial usa informação externa obtida por esses meios como um *input* para a identificação de regras e modelos subjacentes ao confiar em perspectivas, como o aprendizado de máquinas, o qual descreve métodos que auxiliam os computadores a aprenderem sem serem explicitamente programados. Entretanto, a Inteligência Artificial é mais ampla que o próprio aprendizado de máquinas, uma vez que também cobre a habilidade de um sistema de perceber os dados ou de controlar, mover e manipular objetos, com base nas informações aprendidas, seja por meio de um robô, seja por outro dispositivo conectado. Os sistemas de Inteligência Artificial podem ser atualmente classificados como analíticos, inspirados em humanos e humanizados. Os primeiros dispõem de características da Inteligência cognitiva, criam representações do mundo e usam o conhecimento com base em experiências passadas para a tomada de decisões futuras. Já os segundos possuem elementos de inteligência cognitiva e emocional, inclusive emoções na tomada de decisão, enquanto os últimos reúnem competências de inteligência cognitiva emocional e social e são capazes de ter consciência própria nas interações com outros” (STEIBEL, Fabro; VICENTE, Victor; JESUS, Diego Santos Vieira de. Possibilidades e Potenciais da Utilização da Inteligência Artificial. *In*: FRAZÃO, Ana (coord.) et. al. **Inteligência Artificial e Direito: Ética, Regulação e Responsabilidade**. São Paulo: Thomson Reuters Brasil, 2019, p. 55).

⁹⁷ “Art. 35 - São deveres do magistrado:

I - Cumprir e fazer cumprir, com independência, serenidade e exatidão, as disposições legais e os atos de ofício” (BRASIL. Lei Complementar nº 35, de 14 de março de 1979. Dispõe sobre a Lei Orgânica da Magistratura Nacional. **Diário Oficial da União**: Brasília, DF, 14 mar. 1979).

⁹⁸ CONSELHO NACIONAL DE JUSTIÇA. **Resolução nº 60, de 19 de setembro de 2008**. Institui o Código de Ética da Magistratura Nacional. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/127>. Acesso em: 09 dez. 2021.

solução dos casos que lhe sejam submetidos”.⁹⁹ Tal modelo decorre da inafastabilidade da jurisdição (art. 5º, XXXV, da Constituição Federal).¹⁰⁰

Dessa maneira, a figura do juiz de carne e osso, dotado de humanidade, consciência e autoridade¹⁰¹, não pode ser simplesmente substituída ou driblada por um modelo puramente maquinal.

1.4 O Juiz Ciborgue: Riscos e Possibilidades da Interação entre Inteligência Artificial e Teoria da Decisão Judicial

A tutela da decisão judicial por meio da Inteligência Artificial traz uma série promessas de melhorias na prestação do serviço jurisdicional, como: celeridade processual; maior objetividade no julgamento; segurança jurídica no respeito aos precedentes judiciais; respeito à estrutura das regras jurídicas aplicadas; e maior controle de eventuais desvios éticos de magistrados.

Entretanto, podem ser levantados os seguintes riscos inerentes à utilização de Inteligência Artificial na decisão judicial: desumanização das decisões; delegação de atividade jurisdicional à Inteligência Artificial; e preconceitos inerentes ao enviesamento dos algoritmos (*bias*), causado pela base de dados (*dataset*) não representativa da realidade ou por preconceitos do programador/desenvolvedor.

Veja-se que o direito é dotado de complexidade e fluidez que, muitas vezes, escapam à compreensão dos seus próprios operadores. Observa-se que, nos trilhos sinuosos da história, os juristas e filósofos vêm travando intenso debate acerca dos elementos constitutivos do direito, de suas características principais, das fontes de direito legítimo (ou válido) e das adequadas formas de interpretação jurídica. Ainda que atualmente exista um estado da arte do direito ocidental a delinear mínimos denominadores comuns no debate, resta claro que existem campos cujo embate filosófico está longe de ter solução. Todavia, visões positivistas ao estilo de Hans

⁹⁹ Id.

¹⁰⁰ LEAL, Rogério Gesta; SWAROVSKY, Aline. As Garantias Constitucionais do Processo no Brasil. **Revista de Direitos Fundamentais & Democracia**, Curitiba, v. 9, n. 9, p. 242-277, 2011.

¹⁰¹ “A incerteza que permeia o Judiciário não é reduzida ou os custos das decisões não são diminuídos quando outras instâncias assumem o papel de substituir o Judiciário diante de sua autocontenção. Ainda que os tribunais possuam, de fato, limitações e possam tomar decisões equivocadas, o Judiciário é detentor de vantagens institucionais se comparado com outros espaços de tomada de decisão” (LORENZETTO, Bruno Menezes; CLÈVE, Clèmerson Merlin. Teorias Interpretativas, Capacidades Institucionais e Crítica. **Revista Direitos Fundamentais & Democracia**, Curitiba, v. 19, n. 19, p. 164, jan./jun. 2016).

Kelsen¹⁰² e Herbert Hart¹⁰³, e pós-positivistas, nos enfoques de pensadores como Jürgen Habermas¹⁰⁴, Ronald Dworkin¹⁰⁵ ou Robert Alexy¹⁰⁶ demonstram que os modelos filosóficos de interpretação do direito são variados. O direito é um campo de estudo humano, com desavenças igualmente humanas.

Dessa maneira, o modelo a ser pensado deve levar em consideração a interação entre a Inteligência Artificial e a teoria da decisão judicial, desenhada num modelo de juiz independente e em posição imparcial no exercício da indelegável função jurisdicional.

Em que pese a existência de diversos paradigmas de Inteligência Artificial¹⁰⁷, existe uma visão em parte da comunidade jurídica que aponta no sentido de que um modelo de implementação de decisões judiciais com o auxílio de Inteligência Artificial levaria naturalmente à objetividade das decisões, haja vista que o juiz estaria melhor imunizado de uma leitura equivocada do processo causada pelo viés cognitivo do próprio julgador.

Entretanto, essa visão existe justamente pelo fato do Direito ter buscado aproveitar as bases de outra área, a Ciência da Computação. Destaca-se o fato de que os algoritmos e os modelos de IA podem também carregar viés equivocado, o que implica na distorção da representação da realidade (perda de acurácia, de sensibilidade e de precisão¹⁰⁸). Muitas vezes, tal falha nos algoritmos e nos modelos de IA decorre desde a simples imprevisão dos fatos analisados, passando pelo enviesamento causado pelos preconceitos e pela limitada visão de mundo dos responsáveis pela programação, e indo até mesmo no sentido de implementação de viés preconceituoso de maneira deliberada pelo agente humano. Esse fenômeno é chamado de *human bias*¹⁰⁹ (viés humano).

¹⁰² Kelsen, Hans. **Teoria Pura do Direito**. Tradução de João Baptista Machado. São Paulo: Martins Fontes, 1998, p. 388.

¹⁰³ Hart, Herbert L. A. **O Conceito de Direito**. Tradução: Antônio de Oliveira Sette-Câmara. São Paulo: WMF Martins Fontes, 2009, p. 164.

¹⁰⁴ Habermas, Jürgen. **Teoria do Agir Comunicativo**, 1, p. 30-62.

¹⁰⁵ Dworkin, Ronald. **Levando os Direitos a Sério**. 3.^a ed. Tradução: Nelson Boeira. São Paulo: Martins Fontes, 2010, p. 21-127.

¹⁰⁶ Alexy, Robert. **Teoria de Los Derechos Fundamentales**. Madri: Centro de Estudios Constitucionales, 1997, p. 98-147.

¹⁰⁷ O campo da inteligência artificial passa por quatro macrodefinições do que as máquinas podem fazer: a) pensar como um humano; b) agir como um ser humano; c) pensar racionalmente; e d) agir racionalmente (Russell, Stuart; Norvig, Peter. Op. cit., p. 3-4).

¹⁰⁸ Sobre os aspectos técnico do tema, vide o subcapítulo 5.6.

¹⁰⁹ Yao, Adrienne; Weiss, Joseph. Ethical Implications of Bias in Machine Learning. **Proceedings of 51st Hawaii International Conference on System Sciences**, Waikoloa Village, p. 5366, 2018.

Portanto, há uma ideia de que a utilização de Inteligência Artificial surtiria efeitos no sentido de servir como um antídoto contra o viés cognitivo ou contra outros vícios decisórios decorrentes de características inerentes à pessoa do juiz¹¹⁰, pois o magistrado seria alertado acerca de eventuais desvios interpretativos do direito aplicado. Contudo, tal antídoto traz consigo o risco decorrente do viés cognitivo (*human bias*¹¹¹) inerente ao agente responsável pela programação dos algoritmos de Inteligência Artificial. É o velho dilema da decisão judicial e da própria teoria do direito: onde se situa o limite entre a aplicação atada às teias legais e a criação de soluções interpretativas pelo juiz? A Inteligência Artificial potencializou esse debate.

Assim, a Inteligência Artificial aplicada à decisão judicial não pode se desacoplar da consciência¹¹², que pode ser considerada uma unidade de representação¹¹³, a qual somente existirá na figura humana do juiz que percebe cada caso para além da frieza dos números dos *petabytes* de *big data*.¹¹⁴

Nesse sentido, uma máquina de interpretação do Direito dotada de suposta coerência de pensamento filosófico e munida de visão de mundo unificadora seria a

¹¹⁰ “There is no assurance, of course, that algorithms will avoid cognitive biases. They could be built to as to display them. The only point is that they can also be built so as to improve on human decisions, and precisely because they are bias-free. This is simply a specification of the old finding that statistical prediction often outperforms clinical prediction” (SUNSTEIN, Cass R. *Algorithms, Correcting Biases. Social Research: Algorithms*, New York, v. 86, n. 2, 2019, p. 9).

¹¹¹ “Como um sistema probabilístico, baseado numa alimentação de dados para determinada conduta, os algoritmos de *machine learning* podem trazer benefícios ou problemas que refletem discriminação, parcialidades, escolhas ofensivas, desinformações, ou seja, *human bias* ocorre quando o sistema computacional basicamente reflete os valores implícitos de seu criador, distorcendo o conjunto de dados para o treinamento do sistema” (HARTMANN PEIXOTO, Fabiano; SILVA, Roberta Zumblick Martins da. Op. cit., p. 35).

¹¹² HARARI, Yuval Noah. *Homo Deus*, p. 313.

¹¹³ “Denomino também à unidade desta representação, unidade transcendental da consciência, para indicar a possibilidade do conhecimento ‘a priori’ que dali resulta. Porque as diversas representações, dadas em certa intuição, não seriam todas juntas a minha representação, se todas também não pertencessem a uma mesma consciência; quer dizer, que como representações minhas (ainda que não tenha consciência delas como minhas) devem conformar-se necessariamente com a condição, mediante a qual só podem coexistir em uma consciência geral, pois de outro modo não poderiam pertencer-me” (KANT, Immanuel. *Crítica da Razão Pura*. Edição do Kindle, 2020, posição 1812-1817).

¹¹⁴ “O conceito de *Big Data* pode implicar ainda, juntamente com o conceito de *Data Science*, a capacidade de transformar dados brutos em gráficos e tabelas que permitam a compreensão do fenômeno a ser demonstrado. É importante mencionar que, em um contexto em que decisões são tomadas cada vez mais com base em dados, é de extrema importância garantir a veracidade destas informações.

Nas palavras de Maike Wile, *Big Data* é mais que um emaranhado de dados, pois é essencialmente relacional. Apesar de isso não ser um fenômeno novo, o que a internet fez foi dar uma nova dimensão, transformando-o. Para bem entender essas transformações, segundo Wile, precisamos compreender que o *Big Data* somos nós.

A combinação entre objetos inteligentes e *Big Data* poderá alterar significativamente a maneira como vivemos” (MAGRANI, Eduardo. *Entre Dados e Robôs: Ética e Privacidade da Era da Hiperconectividade*. 2.ª ed. Porto Alegre: Arquipélago Editorial, 2019, p. 23-24).

máquina perfeita para eliminação do pluralismo¹¹⁵ inerente à democracia, que é tão caro ao Estado Democrático de Direito. Pouco de *rule of law* sobra numa distopia pura e simples da *rule of machine*.

Justamente para superar o paradigma de resistência à IA¹¹⁶ e lançar voos mais pautados na realidade deve-se ter em vista que cada modelo (admitindo-se a necessária pluralidade de paradigmas) de aplicação de Inteligência Artificial ao direito deve ser desenhado de modo que cada magistrado possa questionar as decisões e fazer com que sua interpretação do direito seja considerada e aplicada aos seus casos. A independência judicial é princípio que deve ser respeitado, sob pena de eliminação da individualidade de cada juiz e de cada cidadão no processo de decisão.¹¹⁷

Todavia, percebe-se que se compreende o dever de respeito aos precedentes dos tribunais (art. 332, art. 489, §1º, IV e V, art. 926 e 927 do Código de Processo Civil) como essencial à observância do princípio da segurança jurídica¹¹⁸ (art. 5º, *caput*, da Constituição Federal) na compreensão de um sistema de *stare decisis* vertical e horizontal¹¹⁹, inibindo-se assim a mácula de decisionismos personalistas do juiz¹²⁰, e esse é um dos usos mais evidentes para Inteligência Artificial no direito.¹²¹ Aliás, pode-se defender fortemente o sistema de precedentes, ainda que a moda à

¹¹⁵ MALISKA, Marcos Augusto. **Fundamentos da Constituição**: abertura, cooperação, integração. Curitiba: Juruá, 2013, p. 55-56.

¹¹⁶ “A resistência para tornar redes e organizações mais eficientes de um modo geral é alta. Uma quantidade grande de esforço será direcionada para superar essa resistência, e o resultado seria uma melhoria anual da capacidade total da humanidade de apenas alguns pontos percentuais. Além disso, mudanças no ambiente interno e externo fazem com que as organizações, mesmo que tenham sido eficientes em algum momento, logo se tornem inadequadas para as novas circunstâncias” (BOSTROM, Nick. Op. cit., posição 2425).

¹¹⁷ “Em vez de indivíduos escolhendo o seu lugar, as máquinas podem se recusar até mesmo a reconhecer que existimos como indivíduos. Somos meras agregações de dados, e à medida que esses dados se fracionam e são distribuídos pelo vasto território da internet, perdemos substância no processo. Quando a nossa geladeira começa a conversar com a luminária da cozinha, sem nenhuma interferência consciente da nossa parte, o que seremos para elas, além de um número de cartão de crédito e um conjunto de atitudes incoerentes? Como define o historiador israelense Yuval Noah Harari, a revolução ameaça ‘desindividualizar’ a todos nós” (RUNCIMAN, David. Op. cit., p. 214).

¹¹⁸ BARBOZA, Estefânia Maria de Queiroz. **Stare decisis, integridade e segurança jurídica**: reflexões críticas a partir da aproximação dos sistemas de common law e civil law. 2010. Tese (Doutorado em Direito). Pontifícia Universidade Católica do Paraná (PUC-PR), Curitiba, 2010, p. 179.

¹¹⁹ DWORKIN, Ronald. **O Império do Direito**. Tradução de Jefferson Luiz Camargo. São Paulo: Martins Fontes, 1999, p. 275-276.

¹²⁰ MACCORMICK, Neil. **Retórica e o Estado de Direito**: Uma teoria da argumentação jurídica. Tradução: Conrado Hübner Mendes e Marcos Paulo Veríssimo. Rio de Janeiro: Elsevier, 2008, p. 339.

¹²¹ CELLA, José Renato Gaziero; WOJCIECHOWSKI, Paola Bianchi. Inteligência Artificial nos Processos Judiciais Eletrônicos. In: FEITOSA, Raymundo Juliano Rego (Org.); et. al. **Direito e novas tecnologias**, Coleção Conpedi/Unicuritiba. Vol.12. São Paulo: Clássica Editora, 2014.

brasileira tenha sido de sabor não tão palatável quando o prometido desde o surgimento das Súmulas Vinculantes (art. 103-A da Constituição Federal e Lei 11.417/2006) até o advento do vigente Código de Processo Civil.

No entanto, uma solução possível seria o treinamento¹²² de cada magistrado em aspectos ligados à tecnologia envolvida na Inteligência Artificial (apenas isso já é um desafio que demanda estudo próprio), a fim de que os juízes estejam aptos a gerenciar um novo tipo de profissional em seus gabinetes (o assessor-desenvolvedor¹²³). Nesse sentido, a acurácia (ou o grau de correção, para utilizar termo jurídico) do juiz (ou do órgão judicante de Tribunal) e de sua assessoria será aprimorada ao invés de ser substituída.¹²⁴

Um Judiciário composto por juízes puramente carimbadores implicaria numa democracia de pior qualidade¹²⁵, pois o magistrado deve sempre ter a possibilidade de, dado o momento histórico, causar uma mudança social ou uma visão jurídica nova, percebendo o momento em que a interpretação deve se adaptar¹²⁶. Nesse sentido, a

¹²² “Os empregados precisarão adaptar constantemente suas funções e habilidades por meio de um processo de aprendizado contínuo, e as máquinas deverão ser monitoradas para que se evitem decisões autônomas que tragam implicações jurídicas e éticas negativas” (STEIBEL, Fabro; VICENTE, Victor; JESUS, Diego Santos Vieira de. Op. cit., p. 62).

¹²³ Sobre a necessária abordagem interdisciplinar para se investigar as interações humano-máquina e suas implicações para as diversas atividades laborativas (inclusive no âmbito das profissões jurídicas), veja-se a Recomendação sobre Ética da Inteligência Artificial elaborada pela UNESCO: “119. Member States should encourage and support researchers to analyse the impact of AI systems on the local labour environment in order to anticipate future trends and challenges. These studies should have an interdisciplinary approach and investigate the impact of AI systems on economic, social and geographic sectors, as well as on human-robot interactions and human-human relationships, in order to advise on reskilling and redeployment best practices” (UNESCO. Op. cit., p. 19).

¹²⁴ Neste aspecto, conforme premissas delineadas no presente estudo, tendo em vista que não é adequada a simples substituição dos operadores do Direito humanos, discorda-se parcialmente da seguinte posição de Rômulo Valentini: “A utilização de sistemas especializados para a elaboração de minutas, em verdade, não substitui a princípio o trabalho dos juízes, mas sim o trabalho de computação humana rotineiramente realizado por assessores e estagiários para elaboração de minutas e realização de estudos e pesquisas jurisprudenciais para subsidiar os fundamentos da decisão” (VALENTINI, Rômulo Soares. **Julgamento por Computadores?** As Novas Possibilidades da Juscibernética no Século XXI e Suas Implicações Para o Futuro do Direito e do Trabalho dos Juristas. 2017. Tese (Doutorado em Direito). Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, 2017, p. 106-107).

¹²⁵ “Capturar os árbitros dá ao governo mais que um escudo. Também oferece uma arma poderosa, permitindo que ele imponha a lei de maneira seletiva, punindo oponentes e favorecendo aliados” (LEVITSKY, Steven; ZIBLAT, Daniel. **Como as Democracias Morrem**. Rio de Janeiro: Zahar, 2018, p. 81-85).

¹²⁶ E aqui, guardado o escopo desta nota de rodapé, modelos positivistas e pós-positivistas encontram pontos de ligação, como, por exemplo, a textura aberta da linguagem utilizada nas normas (que confere um dever interpretativo, segundo numa abordagem pós-positivista, ou discricionariedade judicial, segundo o positivismo) e o papel atribuído aos juízes (de busca de resposta correta inerente à dinamicidade do sistema, para pós-positivistas, ou de atuação suplementar ao legislador voltada à adequação social, para positivistas), ainda que a deferência maior (entre positivistas) ou menor (entre pós-positivistas) ao legislador seja marcante.

concepção de múltiplos significados históricos inerente ao *living constitution*¹²⁷ também abarca a expressão jurídica. A aplicação do direito por meio mecânico tipicamente formalista¹²⁸, ainda que com a possibilidade de adoção de soluções de sistemas preditivos baseados em *machine learning*, causa o risco de que o juiz se acomode ao aplicar de maneira acrítica sugestões de julgamentos dadas por sistemas de Inteligência Artificial. O cachimbo (Inteligência Artificial) tenderá a entortar a boca do usuário do sistema computacional, haja vista a manifestação do “viés do status quo”¹²⁹.

A usabilidade decorrente da transparência (conceito inerente à interface amigável de programas de computador, que esconde os algoritmos que a fazem funcionar)¹³⁰ nos aplicativos poderá causar a falsa sensação de segurança aos juízes. Esse perigo se dá porque uma máscara, na forma de uma interface de sedutora facilidade de utilização, pode levar à ausência de percepção das engrenagens por trás de cada aplicativo de Inteligência Artificial. A passagem do juiz ciborgue ao juiz marionete dos algoritmos é preocupante. Por esse motivo, para superar o receio frente ao novo¹³¹, o desenvolvimento de sistemas de Inteligência Artificial judicial deve incluir

¹²⁷ “Multiple Meanings: This is the view that the constitutional text has multiple linguistic meanings and that constitutional practice should choose between these meanings on a case-by-case basis. Versions of this idea can be found in the work of Professors Richard Fallon and Cass Sunstein” (SOLUM, Lawrence B. Originalism Versus Living Constitutionalism: The Conceptual Structure of the Great Debate. **Northwestern University Law Review**, Chicago, v. 113, n. 6, p. 1243-1296, 2019).

¹²⁸ “Ocurre, por otra parte, que el punto de vista del formalismo dogmático, si es desafortunado al discurrir sobre el papel del intérprete, carece también de la aptitud descriptiva del ser real de la ley que atribuyen sus defensores. Y, por consiguiente, también de eficacia explicativa acerca del modo de producirse efectivamente la actividad del intérprete” (IBÁÑEZ, Perfecto Andrés. ¿Neutralidad o Pluralismo en La Aplicación Del Derecho? Interpretación Judicial e Insuficiencia Del Formalismo. **Doxa**, Madri, n. 15-16, p. 861-882, 1994).

¹²⁹ “Para citar outro exemplo, considere o ‘viés do statu quo’, nome rebuscado para a inércia. Por uma série de razões, que examinaremos mais adiante, as pessoas apresentam uma forte tendência a seguir o statu quo ou a aceitar a opção-padrão” (THALER, Richard H.; SUNSTEIN, Cass R. **Nudge**: como tomar melhores decisões sobre saúde, dinheiro e felicidade. Tradução: Ângelo Lessa. Rio de Janeiro: Objetiva, 2019, posição 180. E-book Kindle).

¹³⁰ “Their transparency is precisely why this class of objects is so appealing to engineers and designers as platforms for computation. These things are already everywhere, hiding in plain sight; nobody bats an eyelash at them. If nothing else, they offer a convenient place to stash the componentry of a computing power that might otherwise read as oppressive” (GREENFIELD, Adam. Op. cit., p. 19).

¹³¹ “Reconheço o argumento ético-jurídico que proponho não exclui de antemão um robô que elabore minutas de decisões; afinal, essas decisões tem um responsável, que é quem as assina e, assim, as profere. O problema é que, diferentemente do recurso e um colaborador ou assistente humano, a quem se confia a minuta, o recurso ao algoritmo cria, já por razões relativas ao volume de minutas que ele é capaz de preparar, uma tentação quase que irresistível de aceitação acrítica do trabalho da máquina como definitivo, que se tornará ainda maior à medida em que a máquina for aperfeiçoando-se. Com isso, o assessor-robô operará como o cavalo de troia que promoverá, de forma inevitável, uma progressiva e velada obsolescência do juiz humano, o único juiz digno desse nome, em favor da máquina” (GRECO, Luís. **Poder de Julgar Sem Responsabilidade de Julgador**: impossibilidade jurídica do juiz-robô. São Paulo: Marcial Pons, 2020, p. 63).

pontos de parada reflexiva previamente estabelecidos¹³², sem uma execução com a pura e simples sugestão da decisão final, a fim de determinar que o julgador conheça todas as opções razoáveis¹³³ de decisões válidas (e de argumentos possíveis¹³⁴) dentro do sistema jurídico.

Ainda na linha do risco de controle dos juízes pela Inteligência Artificial, do ponto de vista político-institucional há perigo em eventuais políticas de controle indireto do teor de decisões judiciais pelo argumento da necessidade de respeito aos precedentes.¹³⁵ Nesse aspecto, cada Tribunal pode “estimular” os magistrados a seguirem seus precedentes mediante as decisões pré-programadas automatizadas em sistemas de computação com processamento centralizado (e este é o modelo até agora mais pensado no Brasil), de maneira que a produtividade individual dos magistrados assim alinhados será sempre “exemplar” e com grande repercussão quantitativa. Já àquele juiz que tenha posição diferente não haverá a possibilidade de automatização de suas decisões, ficando o mesmo à mercê das menores velocidades de processamento e de produção de assessoria humana. Os comparativos quantitativos de produtividade já existem junto ao Conselho Nacional de Justiça¹³⁶ e às Corregedorias dos Tribunais, já a qualidade quase ninguém olha e pouco se percebe.¹³⁷ O juiz usuário do sistema de Inteligência Artificial versus o juiz não alcançado e marginalizado pelo modelo de IA adotado.

¹³² “Humanos também reagem a incentivos, mas sofrem forte influência dos *nudges*. Ao implantar adequadamente tanto os incentivos quanto os *nudges*, aumentamos nossa capacidade de melhorar a vida das pessoas e ajudamos a resolver muitos dos problemas da sociedade. E tudo isso sem impedir a liberdade de escolha de cada indivíduo” (THALER, Richard H.; SUNSTEIN, Cass R. Op. cit., posição 194).

¹³³ RAWLS, John. **Justiça e Democracia**. Tradução: Irene A. Paternot. São Paulo: Martins Fontes, 2000, p. 254-255.

¹³⁴ _____. **Teoria da Argumentação Jurídica: A Teoria do Discurso Racional como Teoria da Justificação Jurídica**. 6.^a ed. Tradução: Zilda Hutchinson Schild Silva. Revisão Técnica e Apresentação: Cláudia Toledo. Rio de Janeiro: Forense, 2021, p. 190-191.

¹³⁵ Respeitar precedentes não retira a possibilidade de *overruling* e a técnica do *distinguishing* não é de simplória aplicação.

¹³⁶ CONSELHO NACIONAL DE JUSTIÇA. **Módulo de Produtividade Mensal do Sistema de Estatísticas do Poder Judiciário**. [2022]. Disponível em: <https://www.cnj.jus.br/capg/mapa.php>. Acesso em: 11 jun. 2022.

¹³⁷ Ao regulamentar o art. 93, II, “b”, da Constituição Federal (“aferição do merecimento conforme o desempenho e pelos critérios objetivos de produtividade e presteza no exercício da jurisdição e pela frequência e aproveitamento em cursos oficiais ou reconhecidos de aperfeiçoamento”), a Resolução 106/2010 do Conselho Nacional de Justiça estabelece o único momento em que há uma avaliação, bastante superficial, do teor das decisões. Esta avaliação se dá no momento da promoção por merecimento dos magistrados, oportunidade em que, dentre diversos aspectos (inclusive elementos quantitativos de “não retenção injustificada de autos além do prazo legal”, “produtividade”, “volume da produção” para cada tipo de ato jurisdicional e “tempo médio do processo na Vara” – art. 3º, III, e art. 4º, II, e art. 6º, II, da Resolução 106/2010 do CNJ), é avaliado o “desempenho (aspecto qualitativo da

Feitas tais considerações, observa-se que a atividade jurisdicional é exercida por magistrados (art. 93 e art. 94 da Constituição Federal), os quais gozam de garantias e prerrogativas necessárias à sua independência e à imparcialidade (art. 95 da Constituição Federal). Na mesma linha, o art. 16 do Código de Processo Civil prevê que a jurisdição “é exercida pelos juízes e pelos tribunais em todo o território nacional...”¹³⁸

Portanto, tem-se em vista na conformação do ordenamento constitucional o sentido de a que atividade jurisdicional se traduz em típica função estatal tida por indelegável a outros agentes públicos ou aos particulares. Nesse sentido, no julgamento do PCA 0000303-58.2011.2.00.0000 o Conselho Nacional de Justiça definiu que o Juiz “exerce o monopólio estatal da jurisdição e determinados atos são por ele indelegáveis para preservar-se tal premissa”¹³⁹.

Por outro lado, contra a necessidade de justificação de decisões judiciais com base em sugestões dadas por sistemas computacionais, eventualmente pode ser lançado o argumento utilitarista de que juízes (e assessores) não teriam correção (acurácia, sensibilidade e precisão¹⁴⁰) no processo de justificação (sem ingressar na discussão do acerto ou do desacerto das decisões), característica que por suposta isonomia isentaria os modelos decisórios baseados em Inteligência Artificial de realizar a justificação das decisões computacionalmente propostas.

Tal tentativa de colonização do direito pela tecnologia soa inconstitucional, pois garantir a justificação da decisão¹⁴¹, para além de possibilitar um processo racional

prestação jurisdicional)” através da “avaliação da qualidade das decisões proferidas” ao se aferir os critérios de “redação”, “clareza”, “objetividade”, “pertinência da doutrina e da jurisprudência” (quando citadas) e “respeito às súmulas do Supremo Tribunal Federal e dos Tribunais Superiores”, na forma do art. 4º, I, e do art. 5º da Resolução 106/2010 do CNJ (CONSELHO NACIONAL DE JUSTIÇA. **Resolução nº 106, de 06 de abril de 2010.** Dispõe sobre os critérios objetivos para aferição do merecimento para promoção de magistrados e acesso aos Tribunais de 2º grau. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/168>. Acesso em: 11 jun. 2022).

¹³⁸ BRASIL. **Lei 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: 11 jun. 2022.

¹³⁹ CONSELHO NACIONAL DE JUSTIÇA. **Procedimento de Controle Administrativo 0000303-58.2011.2.00.0000.** Inconstitucionalidade da completa delegação de atos jurisdicionais a Juízes Leigos. Brasília, DF: Conselho Nacional de Justiça, [2011]. Relator: Cons. Jefferson Luis Kravchychyn, 01 de março de 2011. Disponível em: <https://www.cnj.jus.br/Infojuris/2/Jurisprudencia.seam?jsessionid=1463BB48E76D0A3726156EF32E558E60?jurisprudencialJuris=42662&indiceListaJurisprudencia=6&firstResult=900&tipoPesquisa=BANCO>. Acesso em: 11 jun. 2022.

¹⁴⁰ Vide o subcapítulo 5.6.

¹⁴¹ PIRES, Teresinha Inês Teles. Teorias da Argumentação Jurídica e a Prática Discursiva do Supremo Tribunal Federal: Metodologias Analíticas e Estudo de Casos (AC 4.070/DF e ADI 5526/DF). **Revista Direitos Fundamentais & Democracia**, Curitiba, v. 24 n. 2, p. 102-127, 2019.

pelo julgador humano, fomenta a possibilidade de adequada fundamentação inerente ao caráter justificatório exigido pelo art. 93, IX, da Constituição Federal¹⁴² e à abordagem lógico-dialética determinada pelo art. 489, *caput*, II, e §1º, do Código de Processo Civil¹⁴³.

Ademais, o art. 20, §1º, da Lei 13.709/2018 (Lei Geral de Proteção de Dados) claramente estabelece que o Poder Judiciário (controlador¹⁴⁴) “deverá fornecer, sempre que solicitadas, informações claras e adequadas a respeito dos critérios e dos procedimentos utilizados para a decisão automatizada”. Nesse aspecto, mesmo uma decisão semiautomatizada (com proposta de decisão automatizada por meio de sistema computacional e aceitação racional da mesma pelo juiz) deve respeitar o art. 20, §1º, da Lei 13.709/2018. O juiz ciborgue não escapa de seu dever de justificação racional da decisão.

Dessa forma, em razão do dever de independência nenhum tribunal pode obrigar seus Juízes a aceitarem de maneira acrítica as interpretações propostas mediante Inteligência Artificial. Contudo, o próprio modelo de gestão dos processos

¹⁴² “O art. 93, IX, da CF exige que o acórdão ou decisão sejam fundamentados, ainda que sucintamente, sem determinar, contudo, o exame pormenorizado de cada uma das alegações ou provas, nem que sejam corretos os fundamentos da decisão” (BRASIL. Supremo Tribunal Federal (Plenário). **Agravo de Instrumento 791.292 QO-RG (Tema 339 de Repercussão Geral)**. O dever de fundamentação das decisões judiciais não pressupõe o enfrentamento de todas as teses alegadas pelas partes e nem mesmo a correção dos fundamentos utilizados. Brasília, DF: Supremo Tribunal Federal, [2010].Relator: Min. Gilmar Mendes, 23 de junho de 2010. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=AC&docID=613496>. Acesso em: 11 jun. 2022).

¹⁴³ Código de Processo Civil:

“Art. 489. São elementos essenciais da sentença:

(...)

II - os fundamentos, em que o juiz analisará as questões de fato e de direito;

(...)

§ 1º Não se considera fundamentada qualquer decisão judicial, seja ela interlocutória, sentença ou acórdão, que:

I - se limitar à indicação, à reprodução ou à paráfrase de ato normativo, sem explicar sua relação com a causa ou a questão decidida;

II - empregar conceitos jurídicos indeterminados, sem explicar o motivo concreto de sua incidência no caso;

III - invocar motivos que se prestariam a justificar qualquer outra decisão;

IV - não enfrentar todos os argumentos deduzidos no processo capazes de, em tese, infirmar a conclusão adotada pelo julgador;

V - se limitar a invocar precedente ou enunciado de súmula, sem identificar seus fundamentos determinantes nem demonstrar que o caso sob julgamento se ajusta àqueles fundamentos;

VI - deixar de seguir enunciado de súmula, jurisprudência ou precedente invocado pela parte, sem demonstrar a existência de distinção no caso em julgamento ou a superação do entendimento”.

(BRASIL. **Lei 13.105, de 16 de março de 2015**).

¹⁴⁴ “Art. 5º (...) VI - controlador: pessoa natural ou jurídica, de direito público ou privado, a quem competem as decisões referentes ao tratamento de dados pessoais” (BRASIL. **Lei 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 11 jun. 2022).

pode levar o magistrado questionador a se sentir desestimulado à propositura de nova solução interpretativa, uma vez que desafiar precedente garantido pela IA supostamente infalível seria custoso em tempo, recursos humanos e prestígio (afinal os frios *scores* do CNJ e das Corregedorias não avaliam o grau de acerto ou desacerto de decisões, e nem podem legitimamente fazê-lo na função administrativa, mas são implacáveis no que tange ao aspecto da métrica quantitativa). Faça mais, faça mais rápido! Esse é o mantra que os órgãos de controle passam aos magistrados. Se o juiz não tem acesso às “engrenagens” da Inteligência Artificial fatidicamente será atropelado pela estrutura que garante a aplicação da interpretação contrária, sob pena de arcar com a redução de nota no critério quantitativo utilizado para promoções por merecimento, especialmente no que toca ao “volume da produção” para cada tipo de ato jurisdicional e ao “tempo médio do processo na Vara” (art. 4º, II, e art. 6º, II, da Resolução 106/2010 do CNJ).¹⁴⁵

Já em relação ao aspecto ético na aplicação da Inteligência Artificial à decisão judicial, para se diminuir o enviesamento (*bias*) das propostas de decisão realizadas por meio de sistemas com Inteligência Artificial, principalmente devido ao enviesamento advindo da própria base de dados (*dataset*) utilizada no treinamento da IA (o que pode induzir a decisões de cunho discriminatório¹⁴⁶), são necessárias etapas

¹⁴⁵ CONSELHO NACIONAL DE JUSTIÇA. **Resolução nº 106, de 06 de abril de 2010.**

¹⁴⁶ “Em outubro de 2018, aconteceu a International Conference of Data Protection, em que foi feita uma Declaration on Ethics and Protection in Artificial Intelligence. Dentro de um contexto maior de policy e todas as suas limitações para áreas fronteiriças do conhecimento, a declaração apresenta algumas diretrizes éticas importantes.

A primeira delas é um dever de reduzir preconceitos ou discriminações do desenvolvimento e aplicação de soluções de IA. Isso vai ao encontro das necessidades éticas nas etapas de verificação e validação, acima expostas. Dessa constatação, Polonski (2018), em seu artigo *Mitigating algorithmic bias in predictive justice: 4 design principles for AI fairness*, sugere alguns princípios que vão além de parâmetros tradicionais de avaliação da IA, como acurácia e matriz de representation, que busca uma proteção contra preconceitos injustos e que todos os sujeitos devem ter a mesma oportunidade de ser representados nos dados. Esse princípio deve ser verificado, a priori, buscando um equilíbrio nas populações com rico de sub-representação na fase de treinamento, quando modelos eventualmente tendenciosos sejam submetidos a um novo treino, que sejam utilizados algoritmos para atenuar distorções ou tendências indesejadas metodologicamente. A segunda, protection, impõe aos sistemas de machine learning mecanismos para evitar efeitos injustos a indivíduos vulneráveis. A protection demanda um projeto metodológico abrangente, que permita a prevenção de efeitos negativos em função de gênero, raça, etnia, nacionalidade, orientação sexual, religião, vertentes políticas, entre outros, posto que o sistema treinará a ser testado em referências de dados de uma realidade desigual e repleta de preconceitos. A terceira, stewardship, significa que os sistemas de machine learning têm a responsabilidade ativa pela realização da justiça e o compromisso com a inclusão em todas as etapas do desenvolvimento dos modelos de IA. Por último, a authenticity, ou seja, além da confiança nos dados, deve existir uma autenticidade na visão da própria construção (autenticidade nas previsões da IA), pois serão usadas para suporte de alguma decisão” (HARTMANN PEIXOTO, Fabiano; SILVA, Roberta Zumblick da. Op. cit. p. 39-40).

éticas nos processos de verificação, validação, segurança e controle, neste último caso, deve existir também uma análise humana acerca dos resultados.¹⁴⁷

Assim, é inadequado do ponto de vista do próprio conceito de IA supervisionada, bem como é antiético¹⁴⁸, que se deixe o sistema computacional decidir sem uma espécie de controle humano sobre a acurácia do resultado interpretativo e, mais do que isso, sobre a correção da justificação¹⁴⁹ da decisão judicial computacionalmente proposta.

Dessa forma, é necessário que os modelos de Inteligência Artificial aplicados à decisão judicial sejam submetidos à curadoria¹⁵⁰, do Conselho Nacional de Justiça, do Tribunal, do próprio magistrado responsável pela unidade jurisdicional e da assessoria desse, sob pena de restar subvertido o sistema jurisdicional por meio de uma versão do processo sem juiz.

Assim, é perceptível que as máquinas sabem jogar dados (ou com os dados) com elevados graus de acurácia, sensibilidade e precisão. Muitas vezes as técnicas de Inteligência Artificial utilizadas alcançam o resultado almejado pelos programadores. Mas isso não supera o paradigma desenvolvido na teoria do Direito e na teoria do Estado, pois não existe métrica de decisão correta sem processo de

¹⁴⁷ “Ao escrever sobre o que deve ser prioritário em pesquisas de IA, Russel, Dewey e Tegmark (2015) reforçam alguns aspectos que a pesquisa em IA (que busque trazer benefícios para a humanidade) deve apresentar, fugindo de armadilhas camufladas na tecnologia. Como exemplo de prioridade, deve existir a inclusão de etapas éticas nos processos de verificação, validação, segurança e controle:

1) Verificação: na aferição se o sistema atente aos requisitos formais de construção e a aferição de situações críticas de segurança”. (...)

“ 2) Validação: ao projetar sistemas de maneira adequada, deve-se checar se a solução vai além de atender aos requisitos, mas seja adequada e compatível com um bom comportamento sob o ponto de vista ético”. (...)

“ 3) Segurança: a IA é uma subárea da Ciência da Computação e representa uma das suas fronteiras de inovação. Inicialmente, sob o aspecto da segurança, está a proteção que é necessária para evitar o mau uso da própria IA, crescente em ataques cibernéticos. Há, portanto, esse viés defensivo da IA”. (...)

“ 4) Controle: em áreas consideradas sensíveis, esta etapa ganha ainda mais importância. Os projetos em IA devem prever etapas nas quais o trabalho técnico de controle humano deve ser aplicado. Há também, nesta etapa, o destaque para as oportunidades de arquitetura de protocolos de transmissão entre a atuação autônoma da máquina e o controle humano” (Ibid., p. 37-39).

¹⁴⁸ “Não é possível se falar adequadamente em um projeto de IA sem um projeto ético (que coloque centralmente as discussões sobre ética e valores envolvidos), sem uma reflexão das melhores formas de comunicação e esclarecimento sobre os objetivos e impactos do projeto, sem um sistema de controle (tanto para aferição de benefícios quanto para identificação e gestão de danos)” (Ibid., p. 40).

¹⁴⁹ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 257-260.

¹⁵⁰ “Quais habilidades cognitivas poderiam ser instrumentalmente úteis a um agente depende tanto de seus objetivos finais quanto de sua situação. Um agente que tenha acesso a conselhos de especialistas confiáveis pode ter pouca necessidade de possuir sua própria inteligência e conhecimento. Se inteligência e conhecimento têm um custo, tais como tempo e esforço despendidos na sua aquisição ou o aumento das exigências de armazenamento e processamento, então o agente poderá preferir menos conhecimento e menos inteligência” (BOSTROM, Nick. Op. cit., posição 3853).

justificação, não existe bom nível de acerto (acurácia, sensibilidade e precisão¹⁵¹) da decisão sem um juiz, bem como não existe processo sem julgador humano. Esse é o calcanhar de Aquiles (ou o “parafuso solto”) que marca o problema.

Já é possível se antever que alguns diriam que basta que o juiz, no caso concreto, encontre os melhores fundamentos para aquela decisão que fora tomada pela máquina, afinal de contas, ele é o supervisor humano da máquina. Isso seria típica inversão de valores, na medida em que o juiz seria o assessor da máquina, e não o contrário. É a prova da teoria da Inteligência Artificial como agente autônomo (ainda que sem a consciência de uma IA do tipo geral). Do juiz ciborgue ao juiz marionete num estalar de dedos.

Em tal medida, o sobre-humano, portador de incríveis habilidades e infalível *juiz Hércules* de Dworkin¹⁵², é uma ficção que não pode (e não deve) ser implementada, nem mesmo por uma suposta Inteligência Artificial conforme as premissas aqui analisadas, pois existem outros aspectos inerentes à ciência da computação que não foram aqui abordados, como o fato da interpretação do direito, para além dos desacordos interpretativos relatados pelo próprio Dworkin¹⁵³, se caracterizar como um

¹⁵¹ Os conceitos de acurácia, sensibilidade e precisão se relacionam à Ciência de Dados, tema que será abordado nos subcapítulos 5.3 e 5.5.

¹⁵² “O leitor entenderá agora por que chamei nosso juiz de Hércules. Ele deve construir um esquema de princípios abstratos e concretos que forneça uma justificação coerente a todos os precedentes do direito costumeiro e, na medida em que estes devem ser justificados por princípios, também um esquema que justifique disposições constitucionais e legislativas. Podemos apreender a magnitude de tal empreendimento se distinguirmos, no âmbito do vasto material de decisões jurídicas que Hércules deve justificar, uma ordenação vertical e outra horizontal. A ordenação vertical é fornecida por diferentes estratos de autoridade, isto é, estratos nos quais as decisões oficiais podem ser consideradas como controles das decisões tomadas nos níveis inferiores. Nos Estados Unidos, a natureza gradativa da ordem vertical é evidente. A estrutura constitucional ocupa o mais alto nível, as decisões da Suprema Corte e, talvez, de outros tribunais que interpretam essa estrutura, vêm a seguir. As leis promulgadas pelos diferentes órgãos legislativos ocupam o nível seguinte, e abaixo deste, em níveis diversos, vêm as decisões dos diferentes tribunais que desenvolvem o direito costumeiro. Hércules deve organizar a justificação de princípio em cada um desses níveis, de tal modo que a justificação seja consistente com os princípios que fornecem a justificação dos níveis mais elevados. A ordenação horizontal requer apenas que os princípios que devem justificar uma decisão em um nível devem ser também consistentes com a justificação oferecida para outras decisões do mesmo nível” (DWORKIN, Ronald. **Levando os Direitos a Sério**, p. 182-183).

¹⁵³ “Na interpretação, tudo é muito diferente: os críticos literários ou especialistas em direito constitucional cujos valores sejam excepcionalmente diferentes sob um aspecto pertinente tendem à discordância numa área muito ampla de convicções interpretativas”. (...) “Na interpretação, o holismo não é passivo; é muito ativo”. (...)

“Na interpretação, pelo contrário, as diferenças de objetivos de justificação e de aspirações traduzem-se automaticamente em diferenças de método; os argumentos, longe de estarem protegidos contra essas diferenças, são moldados por elas. Nesse sentido, a convergência parece problemática e, na medida em que ocorre, acidental” (DWORKIN, Ronald. **A Raposa e o Porco-espinho**: Justiça e Valor. Tradução: Marcelo Brandão Cipolla. São Paulo: WMF Martins Fontes, 2014, p. 236-237).

problema IA completo e a explosão combinatorial decorrente da tentativa de sua solução com a tecnologia atual.

Por outro lado, o uso responsável da Ciência da Computação e da Ciência de Dados tem o potencial diminuir a discricionariedade judicial, justamente por aumentar a capacidade de percepção do julgador a respeito do mundo fático-jurídico. O que denota que, sem necessitar de poderes míticos, o juiz humano, auxiliado por IA (o que o caracteriza como o juiz ciborgue), pode fazer aprimorar a sua capacidade de decidir corretamente.

Todavia, é possível pensar em modelos de Inteligência Artificial com capacidade técnica de decisão em campos específicos do direito, sendo que, para se evitar a inconstitucionalidade decorrente da pura e simples substituição do juiz pela máquina munida de inteligência computacional, percebe-se que devem ser desenvolvidos modelos de integração Juiz-IA que levem em conta que cada julgador tem seu próprio âmbito de legitimidade judicante, específica atuação em sua unidade jurisdicional e visão de mundo imersa em experiência jurídico-filosófica.

Outro risco tratado é o referente ao enviesamento dos algoritmos de Inteligência Artificial (decorrente das características da base de dados e influências causadas pelos humanos que desenvolvem os algoritmos), problema que somente pode ser enfrentado por um modelo no qual as fases de verificação, validação, segurança e controle da IA sejam acompanhadas por magistrados, outros operadores do direito e juristas. A *human bias* é um risco à imparcialidade, pois preconceitos e visões de mundo distorcidas do programador (ou do operador do direito) podem contaminar a lógica adotada pela Inteligência Artificial, ainda que pautada em *machine learning* ou *deep learning*.

A racionalidade das escolhas pode ter maior qualidade, tendo em vista que a Inteligência Artificial (caso adequadamente programada e observada) não tem vícios de sentimentos internos que podem levar à superação da lógica racional por uma nova programação psicológica advinda da experiência momentânea do julgador.

Ademais, o controle social estimulado pela Lei 13.709/2018 (LGPD) é importante instrumento de fiscalização do ético e adequado emprego da Inteligência Artificial no âmbito do Poder Judiciário, em especial, no que toca à aplicabilidade de modelos de Inteligência Artificial à decisão judicial.

Por outro ponto de vista, existem incentivos à utilização de sistemas com Inteligência Artificial para o incremento de celeridade processual, o aumento do respeito aos precedentes judiciais (mediante a maior conscientização voltada à observância dos mesmos) e o maior controle (por parte da comunidade jurídica) da correção das decisões judiciais (juízes incoerentes serão denunciados pela própria base de dados).

Nessa linha, a exigência de coerência do juiz para com a visão global de suas decisões será cada vez mais passível de se alcançar, uma vez que uma decisão “fora da curva” ficará registrada.

Por outro lado, embora numa primeira vista exista a concepção de que computadores e sistemas de Inteligência Artificial possam melhorar o nível de objetividade nos julgamentos, a visão sobre o tema não pode ser simplista, pois existem diversas visões do que pode vir a ser considerado um julgamento imparcial.

Ainda no aspecto relativo aos riscos e às possibilidades da interação humano-máquina através de julgamentos auxiliados por Inteligência Artificial, tradicionalmente o tema da imparcialidade desperta o interesse dos estudiosos do direito processual, que se mostraram muito preocupados com a proteção do julgamento em relação a interesses explícitos, anseios encobertos, conflitos de interesses e vieses implícitos.

O debate sobre os elementos objetivos e subjetivos do que se convencionou chamar de princípio da imparcialidade se traduziu na clássica visão de que o juiz se coloca entre as partes e acima delas, sendo que “esta é a primeira condição para que possa exercer sua função dentro do processo”.¹⁵⁴ A esse respeito, com o enfoque em garantir a imparcialidade do julgador, a legislação processual trata dos casos de impedimento e suspeição (art. 144 e art. 145 do CPC).

Além disso, ao tratar da atuação dos magistrados, o art. 1º do Código de Ética da Magistratura prevê que o magistrado deve se nortear pelos princípios da independência e da imparcialidade, dentre outros. O art. 8º do mesmo diploma normativo prevê que o magistrado imparcial é o aquele que se baseia nas provas para alcançar a “verdade dos fatos, com objetividade e fundamento, mantendo ao longo de todo o processo uma distância equivalente das partes, e evita todo o tipo de comportamento que possa refletir favoritismo, predisposição ou preconceito”.¹⁵⁵

¹⁵⁴ CINTRA, Antonio Carlos de Araújo; GRINOVER, Ada Pellegrini; DINAMARCO, Cândido Rangel. **Teoria Geral do Processo**. 19ª ed. São Paulo: Malheiros, 2003, p. 51.

¹⁵⁵ CONSELHO NACIONAL DE JUSTIÇA. **Resolução nº 60, de 19 de setembro de 2008**.

Sobre o tema, o art. 8º, seção 1, da Convenção Americana de Direitos Humanos (1969)¹⁵⁶ prevê que toda pessoa tem direito que o julgamento seja realizado “por um juiz ou Tribunal competente, independente e imparcial, estabelecido anteriormente por lei”, proposta que também consta no art. 10 de Declaração Universal dos Direitos Humanos de 1948.

Dessa forma, sem a pretensão de vencer o tema da imparcialidade (frente aos limites da presente obra), mas com a observação necessária aos riscos e às possibilidades da utilização de IA no auxílio às decisões judiciais, percebe-se que a teoria processual trabalha a imparcialidade como pressuposto para a instauração válida da relação processual através das vedações e garantias à pessoa do Juiz (art. 95 da Constituição Federal), bem como dos aspectos institucionais para a escolha impessoal do julgador e para sua manutenção no processo, mediante direitos fundamentais à proibição de tribunal de exceção e ao juiz natural (art. 5º, XXXVII e LIII, da Constituição Federal).¹⁵⁷

Nesse sentido, o Supremo Tribunal Federal, por exemplo, trata a imparcialidade como uma manifestação objetiva¹⁵⁸, uma vez que a referida garantia processual é inerente ao direito fundamental à isonomia (art. 5º, *caput*, da Constituição Federal), sendo que: a) com base no direito fundamental ao juiz natural, impede o direcionamento do processo a determinado do juiz quando é possível se fazer a distribuição através de sorteio para outros julgadores legalmente competentes para análise do caso¹⁵⁹; b) afasta a possibilidade de que aquele que possui interesse direto

¹⁵⁶ BRASIL. **Decreto 678, de 06 de novembro de 1992**. Convenção Americana de Direitos Humanos. Disponível em: http://www.planalto.gov.br/ccivil_03/decreto/d0678.htm. Acesso em: 28 dez. 2022.

¹⁵⁷ CINTRA, Antonio Carlos de Araújo; GRINOVER, Ada Pellegrini; DINAMARCO, Cândido Rangel. Op. cit., p. 52.

¹⁵⁸ “As causas de impedimento e suspeição do julgador têm por escopo a garantia da imparcialidade mediante a observância dos postulados do juiz natural (artigo 5º, LIII, da CRFB) e do devido processo legal (artigo 5º, LIV, da CRFB), sendo descabida a interpretação ampliativa do artigo 252, III, do CPP que pudesse resultar na criação de situações que permitam à parte interessada escolher quem deixará de examinar sua pretensão, vulnerando-se, por via transversa, os referidos institutos”. (BRASIL. Supremo Tribunal Federal (1. Turma), **Recurso Ordinário em Habeas Corpus 165393 AgR/PE**. Ato alheio à jurisdição do Supremo Tribunal Federal por meio de Recurso Ordinário, bem como incabível a análise do *habeas corpus* (rol taxativo do art. 102, I, “d” e “i”, da Constituição Federal). Brasília, DF: Supremo Tribunal Federal, [2019]. Recorrente: Eudes Teixeira de Carvalho Júnior. Recorrido: Ministério Público Federal. Relator: Min. Luiz Fux, 15 de março de 2019. Disponível em: <https://jurisprudencia.stf.jus.br/pages/search/sjur400729/false>. Acesso em: 10 dez. 2021).

¹⁵⁹ “Distribuição. Direcionamento injustificado da causa a determinado juízo. Ato não aleatório. Ofensa aos princípios do juiz natural e da distribuição livre, que asseguram a imparcialidade do juiz e integram o justo processo da lei. Nulidade processual absoluta. Desnecessidade de indagação de prejuízo. Recurso extraordinário conhecido e provido. Aplicação do art. 5º, XXXVII e LIV, da CF. Distribuição injustificada de causa a determinado juízo ofende o justo processo da lei (due process of law) e, como tal, constitui nulidade processual absoluta” (BRASIL. Supremo Tribunal Federal (2. Turma). **Agravo de**

ou indireto no resultado atue como julgador no processo¹⁶⁰; e, c) levando-se em conta o sistema penal acusatório, impede que o juiz atue de maneira conjunta nas funções de investigar, acusar e julgar.¹⁶¹

Desse modo, na condição de um direito fundamental inerente à igualdade na aplicação da lei¹⁶², o que deve levar em conta as desigualdades naturais entre as partes¹⁶³, a imparcialidade do juiz é uma garantia de justiça para as partes do processo, tendo em vista que o Estado, ao reservar para si a função jurisdicional, “tem o correspondente dever de agir com imparcialidade na solução das causas que lhe são submetidas”.¹⁶⁴

Instrumento 548203 ED/AL. Nulidade absoluta do processo por violação ao princípio do juiz natural. Recorrente: Estado de Alagoas. Brasília, DF: Supremo Tribunal Federal, [2008]. Recorrido: Vanildo de Magalhães Maurício. Relator: Min. Cezar Peluso, 12 de fevereiro de 2008. Disponível em: <https://jurisprudencia.stf.jus.br/pages/search/sjur88959/false>. Acesso em: 10 dez. 2021).

¹⁶⁰ “Ação penal privada. Apelação. Sentença que rejeitou a queixa-crime por falta de preparo. Juiz suspeito. Nulidade. Sentença anulada para dar seguimento à queixa-crime. Apelação prejudicada. 1. Embora o Juiz sentenciante tenha acolhido os fundamentos expostos no parecer do Ministério Público do Estado de São Paulo, as circunstâncias fáticas nas quais estava ele envolvido evidenciam ausência de imparcialidade para julgar causa em que figura como autor pessoa objeto de denúncia oferecida pelo Ministério Público, oriunda de representação criminal, em que figura como ofendido” (BRASIL. Supremo Tribunal Federal (Plenário). **Questão de Ordem no Inquérito 2503/SP**. Nulidade de sentença pela ausência de imparcialidade do juiz. Brasília, DF: Supremo Tribunal Federal, [2008]. Querelante: José Ursílio de Souza e Silva. Investigado: José Abelardo Guimarães Camarinha. Relator: Min. Menezes Direito, 13 de março de 2008. Disponível em: <https://jurisprudencia.stf.jus.br/pages/search/sjur3049/false>. Acesso em: 10 dez. 2021).

¹⁶¹ “Penal e Processual Penal. Imparcialidade judicial e sistema acusatório. Postura ativa e abusiva do julgador no momento de interrogatório de réus colaboradores. Atuação em reforço da tese acusatória, e não limitada ao controle de homologação do acordo. As circunstâncias particulares do presente caso demonstram que o juiz se investiu na função persecutória ainda na fase pré-processual, violando o sistema acusatório. Imparcialidade judicial como base fundamental do processo. Sistema acusatório e separação das funções de investigar, acusar e julgar. Pressuposto para imparcialidade e contraditório efetivos” (BRASIL. Supremo Tribunal Federal (2. Turma). **Recurso Ordinário em Habeas Corpus 144615 AgR/PR**. Inobservância do sistema acusatório no processo em que o juiz tem postura ativa na instrução processual. Brasília, DF: Supremo Tribunal Federal, [2020]. Recorrente: Paulo Roberto Krug. Recorrido: Ministério Público Federal. Relator: Min. Edson Fachin. Relator para acórdão: Min. Gilmar Mendes, 25 de agosto de 2020. Disponível em: <https://jurisprudencia.stf.jus.br/pages/search/sjur435485/false>. Acesso em: 10 dez. 2021).

¹⁶² “Nos detalhes, o dever de igualdade na aplicação da lei apresenta uma estrutura complicada, por exemplo quando exige a elaboração de regras vinculadas ao caso concreto, seja para a precisa determinação de conceitos vagos, ambíguos e valorativamente abertos, seja para o exercício da discricionariedade. No seu núcleo, contudo, esse dever é simples. Ele exige que toda norma jurídica seja aplicada a todos os casos que sejam abrangidos por seu suporte fático, e a nenhum caso que não o seja, o que mais nada significa que dizer que as normas jurídicas devem ser cumpridas” (ALEXY, Robert. **Teoria dos Direitos Fundamentais**. 2ª ed. Tradução: Virgílio Afonso da Silva. São Paulo: Malheiros, 2011, p. 394).

¹⁶³ Ibid., p. 407-415.

¹⁶⁴ CINTRA, Antonio Carlos de Araújo; GRINOVER, Ada Pellegrini; DINAMARCO, Cândido Rangel. Op. cit., p. 52.

Assim, com enfoque nos potenciais de utilização IA judicial, a imparcialidade é avaliável através de elementos objetivos a respeito da posição e do comportamento do julgador no processo.

Ao avançar um pouco mais sobre o tema, pela ótica da teoria da argumentação, julgar de maneira imparcial leva à possibilidade de universalização¹⁶⁵ da solução jurídica adotada, uma vez que quem “pretende tratar uma pessoa *A* de maneira diferente da pessoa *B* está obrigado a fundamentá-lo”¹⁶⁶. Isso ocorre porque, por coerência, quem “afirma ‘ $p \rightarrow q$ ’ deve, se seu interlocutor formula ‘ $\neg q$ ’, ou aceitar ‘ $\neg p$ ’ ou renunciar a ‘ $p \rightarrow q$ ’”.¹⁶⁷ Assim, conforme ressalta Alexy, para “a fundamentação de uma decisão jurídica deve-se apresentar pelo menos uma norma universal” (J.2.1)¹⁶⁸. Em suma, a universalidade atua como base para a imparcialidade, na medida em que a alteridade no tratamento dos diversos casos pode ser demonstrada por meio de uma razão comum, norteadada pela pretensão de correção¹⁶⁹, o que desestimula solipsismos do julgador.

Dessa forma, a ideia de universalização como elemento da imparcialidade aplicada aos sistemas de Inteligência Artificial judicial não se confunde com uma visão de espectador imparcial¹⁷⁰, uma vez que devem ser observadas as influências dialógicas a) quando “não houver razão suficiente para a permissibilidade de um tratamento igual, então, é obrigatório o tratamento desigual”¹⁷¹; e b) se “houver uma razão suficiente para o dever de um tratamento desigual, então, o tratamento desigual é obrigatório”¹⁷². Essas possibilidades de tratamento desigual para situações naturalmente desiguais (isonomia) muitas vezes estão expressas na legislação e nos precedentes. Isso ocorre, por exemplo, quando a legislação processual determina que o fornecedor deve produzir a prova, sob pena da afirmação da parte consumidora ser

¹⁶⁵ “Aja de tal modo que a máxima de sua vontade possa valer ao mesmo tempo como princípio de uma legislação universal” (KANT, Immanuel. **Crítica da Razão Prática**. Tradução: Monique Hulshof. Petrópolis: Editora Vozes; Bragança Paulista: Editora Universitária São Francisco, 2017, p. 42).

¹⁶⁶ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 174.

¹⁶⁷ Id.

¹⁶⁸ Ibid., p. 258.

¹⁶⁹ “Sistemas normativos que não formulam explícita ou implicitamente uma pretensão à correção não são sistemas jurídicos” (ALEXY, Robert. **Conceito e Validade do Direito**. Tradução: Gercélia Batista de Oliveira Mendes. São Paulo: WMF Martins Fontes, 2009, p. 152).

¹⁷⁰ “A argumentação pública é claramente uma característica essencial da objetividade nas crenças políticas e éticas. Se Rawls oferece uma maneira de pensar a objetividade na avaliação da justiça, Adam Smith oferece outra ao invocar o espectador imparcial” (SEN, Amartya. **A Ideia de Justiça**. Tradução: Denise Bottmann e Ricardo Doninelli Mendes. São Paulo: Companhia das Letras, 2009, posição 1136. E-book Kindle).

¹⁷¹ ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 409.

¹⁷² Ibid. p. 410.

tida como verossimilhante (art. 6º, VIII, do Código de Defesa do Consumidor), quando há a distribuição dinâmica do ônus probatório conforme a maior facilidade da respectiva parte produzir a prova (art. 373, §1º, do Código de Processo Civil) ou quando é adotada a interpretação mais favorável ao trabalhador, na dúvida a respeito de cláusula do contrato de trabalho.¹⁷³

Portanto, dentro de um contexto de arena deliberativa baseada na liberdade de comunicação¹⁷⁴ é possível uma objetividade abrangente¹⁷⁵ no processo da era da Inteligência Artificial judicial, a qual tem relação com a equidistância psicológica do julgador em relação às partes e aos seus anseios.¹⁷⁶ Nesse ponto, para casos menos sensíveis¹⁷⁷, a utilização de sistema de IA judicial efetivamente traz ganhos de performance através da maior objetividade no julgamento.

A esse respeito a abordagem de julgamentos através de técnicas de Inteligência Artificial judicial se baseia na maior objetividade consolidada sob o projeto da Teoria do Direito.

O problema é a escolha. Até onde a sociedade pretende ir na integração dos sistemas de Inteligência Artificial e Direito, e isso depende das possibilidades futuras

¹⁷³ “No Direito do Trabalho, se uma cláusula contratual ou normativa é dúbia ou admite mais de uma interpretação, a correta exegese é aquela que permite sanar a dúvida da forma mais favorável ao trabalhador. Trata-se do princípio do ‘in dubio pro operario’” (BRASIL. Tribunal Regional do Trabalho da 1ª Região (10. Turma). **Recurso Ordinário 0010600-02.2015.5.01.0261/RJ**. Interpretação *in dubio pro operario* de cláusula normativa. Rio de Janeiro, RJ: Tribunal Regional do Trabalho da 1ª Região, [2017]. Relator: Des. Flavio Ernesto Rodrigues Silva, 26 de abril de 2017. Disponível em: <https://bibliotecadigital.trt1.jus.br/jspui/handle/1001/1343580?mode=full>. Acesso em: 12 dez. 2021).

¹⁷⁴ “Liberdade comunicativa só existe entre atores que desejam entender-se entre si sobre algo num enfoque performativo e que contam com tomadas de posição perante pretensões de validade reciprocamente levantadas. Essa característica peculiar da liberdade comunicativa, que depende sempre de uma relação intersubjetiva, explica por que ela se liga a obrigações ilocucionárias. Para alguém poder tomar uma posição, dizendo ‘sim’ ou ‘não’, é preciso que o outro esteja disposto a fundamentar, caso se torne necessário, uma pretensão levantada através de atos de fala. Uma vez que os sujeitos que agem comunicativamente se dispõem a ligar a coordenação de seus planos de ação a um consentimento apoiado nas tomadas de posição recíprocas em relação a pretensões de validade e no reconhecimento dessas pretensões, somente contam os argumentos que podem ser aceitos *em comum* pelos partidos participantes. São respectivamente os *mesmos* argumentos que têm uma força racionalmente motivadora” (HABERMAS, Jürgen. **Direito e Democracia**, p. 156).

¹⁷⁵ “Apesar das diferenças entre os distintos tipos de argumentos apresentados por Smith, Habermas e Rawls, há uma semelhança essencial em suas respectivas abordagens da objetividade: cada um deles associa a objetividade, direta ou indiretamente, à possibilidade de sobreviver aos desafios da análise informada proveniente de direções diversas” (SEN, Amartya. Op. cit., posição 1151).

¹⁷⁶ “O liberalismo político vê sua forma de filosofia política como possuidora de um tema próprio, qual seja: como é possível a existência de uma sociedade justa e livre em condições de profundo conflito doutrinário, sem perspectiva de resolução? Para manter a imparcialidade entre doutrinas abrangentes, o liberalismo político não discute especificamente os tópicos morais que dividem essas doutrinas” (RAWLS, John. **O Liberalismo Político**. 2ª ed. Tradução: Dinah de Abreu Azevedo São Paulo: Ática, 2000, p. 36).

¹⁷⁷ Casos repetitivos ou de aplicação de regras para situações jurídicas mais simples, por exemplo.

de desenvolvimento de algoritmos éticos que garantam que os humanos estejam sempre no controle. Tal questão passa pela necessidade de desenvolvimento de algoritmos como espécies de vetores e freios tecnomorais a serem observados pela Inteligência Artificial no momento da tomada de decisões e para que a IA reconheça cada ser humano dentro de sua individualidade.

1.5 Entropia e Direito: P vs. NP, ou, o Mundo é Maior que o Direito

Esta é uma pesquisa sobre Direito, o que pode causar o estranhamento sobre o motivo de se mencionar um dos maiores problemas não resolvidos da Ciência da Computação e da Matemática. Todavia, a questão tratada neste tópico impacta a aplicação do Direito, embora não tenha sido pensada originalmente para o mundo da Filosofia do Direito e da Teoria do Direito.

O problema P vs. NP é um dos “Problemas do Milênio” pendentes de resolução, conforme referido pelo *Clay Mathematics Institute*.¹⁷⁸ A solução do problema provavelmente garantirá, ao pesquisador que assim o consiga, o prêmio de um milhão de dólares e uma medalha *Fields*. Percebe-se que não será a partir das ciências humanas e, muito menos do Direito, que partirá uma solução para o problema P vs. NP.

Todavia, o problema P vs. NP, onde P representa os algoritmos cujo tempo de execução é polinomial e o NP representa as soluções de tempo de execução não polinomial, ao não ter ainda sua equivalência ($P = NP$), provada ou refutada, tem relevância quando se pensa em Direito aplicado com o auxílio de Ciência da Computação.

Por exemplo, ao se pensar na estruturação de dados numéricos, existem diversos algoritmos polinomiais (P) que podem ser utilizados para organizar e estruturar dados. Alguns exemplos dos algoritmos de ordenação de dados mais conhecidos¹⁷⁹ são *insertion sort*, *selection sort*, *bubble sort*, *quick sort*, *merge sort* e *shell sort*.¹⁸⁰

¹⁷⁸ CLAY MATHEMATICS INSTITUTE. **P vs NP Problem**. [2020]. Disponível em: <https://www.claymath.org/millennium-problems/p-vs-np-problem>. Acesso em: 11 jan. 2022.

¹⁷⁹ GANMOTE, Shubham V.; VISHWAS G. R. S.; KUMAR, Anupama. Analysis of Sorting Algorithms Using Time Complexity. **International Journal of Engineering Research & Technology (IJERT)**, Gandhinagar, v. 5, n. 21, 2017.

¹⁸⁰ Tendo em vista que o objetivo da pesquisa não é discutir algoritmos específicos, para uma noção mais lúdica do funcionamento dos algoritmos de ordenação de dados pode ser acessada a seguinte

Em passo mais profundo, percebe-se que problemas do Direito são, no seu cerne, problemas linguísticos.¹⁸¹ Dessa forma a polissemia intrínseca às palavras é uma barreira natural à pretensão de solução computacional dos problemas jurídicos em tempo polinomial (P). Ou seja, ainda que fosse possível catalogar todas as possíveis soluções para uma dada aplicação de uma norma jurídica (a facticidade em interação com a normatividade), o número de ações computacionais (cálculos) seria muitíssimo maior do que o número de elementos fático-jurídicos computados.

Nesse sentido, para além do debate não resolvido de um dos maiores problemas pendentes da Ciência da Computação ($P = NP?$), observa-se que os problemas de aplicação do Direito são resolvidos por soluções não polinomiais (NP).¹⁸² Isso ocorre porque o caminho interpretativo pode passar por vários tipos diferentes de análises fáticas e de argumentos, o que gera um crescimento de ordem exponencial do tempo de execução para se verificar previamente cada caminho interpretativo possível na construção da solução jurídica.

Em suma, separar o problema maior (cujo tempo de execução é NP) em problemas menores, mediante abstrações e tomadas de posições heurísticas, é obrigatório para se pensar em computação de problemas jurídicos.

Em palavras mais jurídicas (e menos matemáticas), dada a complexidade natural dos problemas jurídicos, estes são insondáveis através de soluções puramente matemáticas.¹⁸³ Não se verifica a possibilidade de desenvolvimento de algoritmos que busquem esgotar as possibilidades de soluções de um problema jurídico. Por exemplo, se nem mesmo um jurista experiente consegue antecipar todas as formas de aplicação do direito constitucional à dignidade humana e toda sua rica gama de manifestações, não será um sistema de Inteligência Artificial que o fará por

página disponibilizada no website da Universidade de São Francisco (USFCA): <https://www.cs.usfca.edu/~galles/visualization/ComparisonSort.html> (GALLES, David. **Comparison Sorting Algorithms**. [2020]. Disponível em: <https://www.cs.usfca.edu/~galles/visualization/ComparisonSort.html>. Acesso em: 06 jan. 2022).

¹⁸¹ HART, Herbert A. L. Op. cit., p. 191-192; WARAT, Luis Alberto. **O Direito e Sua Linguagem**. Porto Alegre: Sérgio Fabris, 1984, p. 64.

¹⁸² Por exemplo: "O problema do caixeiro-viajante (TSP) é um problema de roteiro de viagem em que cada cidade deve ser visitada exatamente uma vez. O objetivo é encontrar o percurso *mais curto*. O problema é conhecido por ser NP-difícil, mas um grande esforço tem sido empregado para melhorar os recursos de algoritmos de TSP. Além de planejar viagens para caixeiros-viajantes, esses algoritmos são usados para tarefas como planejar movimentos de máquinas automáticas para perfuração de placas de circuitos e de máquinas industriais em fábricas" (RUSSELL, Stuart; NORVIG, Peter. Op. cit., p. 67).

¹⁸³ A jurimetria, por exemplo, tem sua utilidade para o auxílio à tomada de decisão, mas não indica soluções completas, haja vista se tratar de uma sondagem estatística de como o Direito se manifesta.

meio de uma solução em tempo de execução polinomial (P). Em resumo, segundo a tecnologia hoje existente e o atual estado da Ciência da Computação¹⁸⁴, os problemas jurídicos somente podem ser auxiliados, parcialmente, por meio de computação através de algoritmos não polinomiais (NP).

Dessa maneira, aplicar o Direito com o auxílio de Inteligência Artificial continua a ter, de certa forma, os mesmos desafios que sempre existiram, na medida em que a interpretação jurídica da era da IA ainda é (e por muito tempo será) uma sondagem provisória e parcial de uma realidade jurídica que é mutante e dinâmica.

O cérebro humano adota atalhos para resolver problemas cujas soluções são em tempo de execução não polinomial, isso ocorre através de heurísticas.¹⁸⁵ Nesse sentido, o caráter heurístico do uso não-justificativo de enunciados dogmáticos¹⁸⁶ é uma regra implícita na argumentação jurídica (e na aplicação do direito), pois nenhum “dogmático pode fundamentar todos os enunciados dogmáticos em que se apoia para a discussão de seu problema”¹⁸⁷, uma vez que isso levaria a um processo sem fim de fundamentação. Tal característica suscita o conceito de “parada ótima”¹⁸⁸, que representa o necessário ponto de parada no processo argumentativo, com a escolha da melhor solução conhecida mediante a pretensão de correção, em detrimento de outras soluções desconhecidas (ainda que eventualmente sejam melhores).

Assim, a interpretação do Direito pode ser classificada como um problema NP-completo, pois a complexidade na busca de encontrar e validar computacionalmente

¹⁸⁴ O futuro desenvolvimento da computação quântica pode mudar isso, caso a promessa de uma nova era de poder de processamento se confirme. Mas inexiste certeza disso neste momento.

¹⁸⁵ “A principal função do Sistema 1 é manter e atualizar um modelo de seu mundo pessoal, que representa o que há de normal nele. O modelo é construído por associações que ligam ideias de circunstâncias, eventos, ações e resultados que coocorrem com alguma regularidade, seja ao mesmo tempo, seja dentro de um intervalo relativamente curto” (KAHNEMAN, Daniel. **Rápido e Devagar**: Duas formas de pensar. Tradução: Cássio de Arantes Leite. Rio de Janeiro: Objetiva, 2012, posição 1502. E-book Kindle).

¹⁸⁶ “Portanto, é possível conhecer o conteúdo de um sistema jurídico sem conhecer a identidade de nenhuma das leis. Isto é pressuposto por Kelsen quando diz: ‘É tarefa da ciência jurídica representar o direito de uma comunidade, isto é, o material produzido pela autoridade jurídica no procedimento de criação das leis, na forma de enunciados’ que tenham estrutura. É possível identificar e entender esse material jurídico sem saber como dividi-lo em leis” (RAZ, Joseph. **O Conceito de Sistema Jurídico**: Uma Introdução à Teoria dos Sistemas Jurídicos. Tradução: Maria Cecília Almeida. São Paulo: WMF Martins Fontes, 2012, p. 97).

¹⁸⁷ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 228.

¹⁸⁸ “De modo intuitivo, pensamos que uma tomada de decisão racional tem a ver com enumerar de forma exaustiva nossas opções, avaliar cada uma com cuidado, e depois escolher a que for melhor. Mas na prática, quando o relógio — ou o cronômetro — está tiquetaqueando, poucos aspectos da tomada de decisão (ou, mais genericamente, do processo de pensar) são tão importantes quanto este: quando parar” (CHRISTIAN, Brian; GRIFFITHS, Tom. **Algoritmos para viver**. Tradução: Paulo Geiger. São Paulo: Companhia das Letras, 2016, posição 733-738. E-book Kindle).

todas as soluções possíveis (o tempo e execução) para a interpretação cresce numa escala fatorial, o que representa um desafio que não pode ser enfrentado numa abordagem de completude de computação jurídica.

O exemplo clássico de um problema NP-completo é o caso do caixeiro viajante¹⁸⁹ que deve estabelecer a rota mais curta (mais eficiente) entre diversas cidades a serem visitadas. Nesse exemplo, a cada nova cidade incluída entre as possibilidades surge uma série de novas rotas devem ser verificadas.

Em certa medida, testar novas rotas entre cidades diferentes se assemelha a testar novas possibilidades de interpretação, como novos sentidos para um termo polissêmico, verificação dos critérios fixados em precedentes de diversos tribunais, investigação de posicionamentos doutrinários (argumentos dogmáticos), exercício de técnicas de ponderação, etc. Em muitos casos não existe uma única forma de interpretar o Direito, ainda que se busque a resposta correta.

Por isso, na impossibilidade de busca ao infinito (e de argumentação ao infinito), o cérebro humano adota abordagens de inferência de respostas através de razão prática¹⁹⁰, o que deve ser levado em conta no desenvolvimento de sistemas de Inteligência Artificial utilizados no auxílio à decisão judicial. A partir dessa realidade, observa-se que o Direito da era da Inteligência Artificial continuará a ser indelevelmente humano na forma de interpretação.¹⁹¹

1.6 Da Entropia Jurídica ao Duplo Canal de Correção: Direito e Teoria da Informação

Desde o Século V a. C. o Partenon está assentado na acrópole de Atenas. A bela construção é um tributo à inventividade e à cultura humana, bem como uma vítima de duas forças (senão uma, em verdade) inexpugnáveis, o tempo e a história.

¹⁸⁹ RUSSELL, Stuart; NORVIG, Peter. Op. cit., p. 67.

¹⁹⁰ “Segundo essa teoria, uma declaração normativa é correta ou – sob o pressuposto de uma teoria da verdade liberal – verdadeira, quando ela pode ser resultado de um determinado procedimento, o do discurso racional” (ALEXY, Robert. ALEXY, Robert. **Direito, Razão, Discurso**: Estudos para a filosofia do direito. Tradução: Luís Afonso Heck. 2ª ed. Porto Alegre: Livraria do Advogado, 2015, p. 78).

¹⁹¹ E neste aspecto concorda-se com Luís Greco, quando o referido autor suscita diversas características do que seria uma “boa decisão” judicial (GRECO, Luís. Op. cit., p. 18-20).

A flecha do tempo é implacável, tudo deteriorando e destruindo, e a história é um componente natural às mudanças pelas quais tudo que existe tem de passar.¹⁹²

Assim, o templo de Atena sucumbiu ao tempo e à história (a conquista Romana e o cristianismo têm papéis importantes nisso), e hoje o que restou do Partenon é uma parcela do que um dia o magnífico edifício já significou em aspectos arquitetônicos e culturais.

Essa constatação se deve ao fato de que a força mais inexpugnável do Universo tudo toca, tudo desconstrói. Trata-se da entropia.

A entropia tem diferentes facetas, exploradas por diversos pesquisadores durante o caminho de seu estudo. Desse modo, a partir da Segunda Lei da termodinâmica, a entropia pode ser resumida como a irreversibilidade do processo cumulativo e aleatório de desorganização do mundo, tido como um sistema. Assim, tudo que conta com algum grau de organização sofre a influência da entropia.

A entropia explica fenômenos físicos, informacionais e sociais.¹⁹³ A esse respeito, o caminho pelo qual os estudos sobre a entropia percorreram passa pelos aspectos do funcionamento do Universo até a Teoria da Informação. Nesse sentido, podem ser destacadas as pesquisas de Sadi Carnot¹⁹⁴, Rudolf Clausius¹⁹⁵, Ludwig Boltzmann¹⁹⁶, Josiah Gibbs¹⁹⁷, James Maxwell¹⁹⁸, Max Planck¹⁹⁹ e Claude Shannon.²⁰⁰

¹⁹² “A física de não equilíbrio estuda os processos dissipativos, caracterizados por um tempo unidirecional, e, com isso confere uma significação à irreversibilidade” (PRIGOGINE, Ilya. **O Fim das Certezas**: Tempo, Caos e as Leis da Natureza. Tradução: Roberto Leal Ferreira. 2ª ed. São Paulo: Editora UNESP, 2011, p. 11).

¹⁹³ KAFRI, Oded; KAFRI, Hava. **Entropy**: God’s Dice Game. Scotts Valley: CreateSpace, 2013, posição 2720-2798. E-book Kindle.

¹⁹⁴ Ibid., posição 206.

¹⁹⁵ ROVELLI, Carlo. **A Ordem do Tempo**. Tradução: Silvana Cobucci. Rio de Janeiro: Objetiva, 2017, p. 19.

¹⁹⁶ KAFRI, Oded; KAFRI, Hava. Op. cit., posição 659-957.

¹⁹⁷ Ibid., posição 990-1055.

¹⁹⁸ Ibid., posição 1163-1331.

¹⁹⁹ Ibid., posição 1556-1706.

²⁰⁰ SHANNON, Claude E.; WEAVER, Warren. **The Mathematical Theory of Communication**. Chicago: University of Illinois Press, 1963, p. 13. E-book Kindle.

Dessa forma, conforme expõem Stephen Hawking²⁰¹ e Ilya Prigogine²⁰², a flecha do tempo da desorganização termodinâmica²⁰³ é a mesma flecha do tempo psicológica, motivo pelo qual a percepção de passado e presente se verifica possível, fenômeno que é necessário para o processo comunicacional e sua entropia de dados. Foi Ludwig Boltzmann²⁰⁴ que, na busca de uma idealização teórica que interpretasse a realidade, veio a se deparar com o caráter irreversível e indeterminável do Universo, a entropia que impulsiona a flecha do tempo na percepção humana. De acordo com os estudos do físico austríaco, a flecha do tempo demonstra o fato de que nós, “humanos, observadores limitados, seríamos responsáveis pela diferença entre passado e futuro”.²⁰⁵

Especificamente sobre a relação entre tempo e Direito, pode-se citar os estudos de Cristiano Paixão Araujo Pinto, que fez uma compilação exaustiva da evolução da ciência na busca de explicar o conceito de tempo²⁰⁶, o qual destaca que, para além da consagração da irreversibilidade, a flecha do tempo “representa, muito além disso, um fator de ligação, uma conexão entre o homem e a natureza”.²⁰⁷ Esse fator impede a total positivação jurídica em um núcleo imodificável²⁰⁸, uma vez que o Direito possui

²⁰¹ “Resumindo, as leis da ciência não fazem distinção entre as direções para frente e para trás do tempo. Contudo, há pelo menos três setas do tempo que distinguem o passado do futuro. São elas: a seta termodinâmica, a direção do tempo em que a desordem aumenta; a seta psicológica, a direção do tempo em que nos lembramos do passado, e não do futuro; e a seta cosmológica, a direção do tempo em que o universo se expande, em vez de se contrair. Como já mostrei, a seta psicológica é em essência a mesma que a seta termodinâmica, de modo que as duas sempre apontariam na mesma direção” (HAWKING, Stephen. **Uma Breve História do Tempo**. Tradução: Cássio de Arantes Leite. Rio de Janeiro: Intrínseca, 2015, p. 189).

²⁰² A flecha do tempo é um conceito ligado à percepção humana, conforme destaca Ilya Prigogine: “Sabe-se que Einstein afirmou muitas vezes que ‘o tempo é ilusão’. E, de fato, o tempo real como foi incorporado nas leis fundamentais da física, da dinâmica clássica newtoniana até a relatividade e a física quântica não autoriza nenhuma distinção entre o passado e o futuro. Ainda hoje, para muitos físicos, esta é uma verdadeira profissão de fé: em termos de descrição fundamental da natureza, não há flecha do tempo” (PRIGOGINE, Ilya. Op. cit., p. 10).

²⁰³ “O crescimento da entropia designa, pois, a direção do futuro, quer no nível de um sistema local, quer no nível do universo como um todo” (Ibid., p. 26).

²⁰⁴ “O paradoxo do tempo só foi identificado tardiamente, na segunda metade do Século XIX, graças aos trabalhos do físico vienense Ludwig Boltzmann. Ele acreditava poder seguir o exemplo de Charles Darwin na biologia e fornecer uma descrição evolucionista dos fenômenos físicos. Sua tentativa teve como efeito pôr em evidência a contradição entre as leis da física newtoniana baseadas na equivalência entre passado e futuro, e toda tentativa de formulação evolucionista que afirma uma distinção essencial entre passado e futuro. Na época, as leis da física newtoniana eram aceitas como expressão de um conhecimento ideal, objetivo e completo. Já que as leis afirmavam a equivalência entre o passado e o futuro, toda tentativa de conferir uma significação fundamental à flecha do tempo aparecia como uma ameaça a esse ideal. A situação não mudou hoje” (Ibid., p. 10).

²⁰⁵ Ibid., p. 11.

²⁰⁶ PINTO, Cristiano Paixão Araujo. **Modernidade, Tempo e Direito**. Belo Horizonte: Del Rey, 2002, p. 9-115.

²⁰⁷ Ibid., p. 113.

²⁰⁸ Ibid., p. 248.

complexidade e indeterminação entrópicas que o deixam aberto para o futuro.²⁰⁹ Também pode ser referido o estudo de Ricardo Campos, no qual aponta o tempo como um fator de ordenação do Direito²¹⁰ e perene reorganização jurídica.²¹¹

Com base nessas premissas, dados organizados se transformam em informação. Informação contextualizada se transforma em conhecimento. Conhecimento pode se tornar sabedoria. Assim, o Direito não se resume aos dados fático-jurídicos, mas também não pode desprezar os dados na busca de sua aplicação. Isso insere a entropia no contexto da interpretação jurídica²¹² e dá destaque à possibilidade de utilização de sistemas computacionais para se enfrentar a complexidade. Nesse sentido, frente à complexidade entrópica do mundo do Direito, a Inteligência Artificial jurídica é uma aliada.

²⁰⁹ “Entende-se delineada, nesses termos, a dimensão temporal do direito da sociedade moderna. Futuro em aberto, seletividade elevada, indeterminação, complexidade e contingência são expressões-chave deste fenômeno, segundo observado nas linhas antecedentes.

Demonstrou-se, também, a evolução da dimensão temporal ao longo das formas da diferenciação do direito. Foi possível verificar a crescente importância do papel atribuído ao futuro no direito da sociedade, tendo sido evidenciada, igualmente, a conexão entre esta gradativa abertura do futuro e o processo de diferenciação e autonomização do sistema jurídico” (Ibid., p. 254).

²¹⁰ “O fato de que o futuro esteja à disposição na sociedade moderna nada diz sobre a sua forma de lidar com ele. O tempo em si não pode ser fixado, mas a sociedade por sua vez cria estruturas ou formas de estruturas (de expectativa) para lidar com o um futuro aberto” (CAMPOS, Ricardo. **Metamorfoses do Direito Global**: sobre a interação entre direito, tempo e tecnologia. São Paulo: Contracorrente, 2022, p. 122).

²¹¹ “Essa incerteza exige cada vez mais estruturas sociais para lidar com ela, mas sem ter que, entretanto, abolí-la ou evitá-la completamente. Especialmente para o Direito moderno, a formação de estruturas de expectativa é de importância central neste contexto. Neste sentido, as categorias de tempo, estrutura e expectativa exibem profundas interdependências históricas e técnicas, por um lado, e mostram uma extrema abertura para o futuro, por outro” (Ibid., p. 124-125).

²¹² “Também a relatividade com vista aos participantes cria incerteza. Além disso, ela leva, em todo o caso, em discursos reais temporalmente limitados, a um alargamento do espaço do só discursivamente possível. E, mais além, deve ainda ser considerado que ela é necessária em virtude da estrutura de problemas práticos. Essa exige, como já exposto, um procedimento não-monológico. Acresce a isso, que cada discurso deve ter um ponto de partida. Ele não pode iniciar com nada. Esse ponto de partida consiste nas convicções normativas, cada vez, faticamente existentes, dos participantes. O procedimento do discurso é nada mais que um procedimento de seu estudo minucioso racional. Nisso, cada convicção normativamente relevante é um candidato para uma modificação baseada em argumentação racional. Nessa limitação à estruturação racional do da argumentação reside um mérito importante da teoria do discurso. Uma teoria que tenta não só estruturar racionalmente o processo de argumentação ou decisão, mas, por exemplo, pelo fato de ela prescrever determinados conteúdos como premissas de partida, também determinar racionalmente, está exposta não só a objeções contra as premissas de partida quanto ao conteúdo, selecionadas pelo respectivo teórico da argumentação ou da decisão, às quais regularmente é mais difícil de replicar do que as objeções contra regras do discurso, no essencial, formais, mas, mais além, também à objeção fundamental, que o teórico, com isso, entra em um campo que, já porque suas convicções normativas quanto ao conteúdo não são geralmente mais corretas que as dos participantes, melhor ficar a cargo destes, uma que o papel de participante também ao teórico está aberto a qualquer hora” (ALEXY, Robert. **Direito, Razão, Discurso**, p. 99).

Por exemplo, uma célula, seu núcleo e suas organelas podem ser medidos e descritos com precisão em dados. Um átomo pode ser, no limite da tecnologia atual, dividido em partes cada vez menores (*quarks*), o que pode ser traduzido em dados. Contudo, o conhecimento tem limites, pois o acesso à informação é limitado pela realidade, e isso impacta o Direito.

Numa primeira vista pode ocorrer a errônea compreensão de que o Direito é uma coisa tangível, palpável e aprisionável²¹³ em experimentos mentais.²¹⁴ Isso é um engano, pois a totalidade dos dados não pode ser integralmente conhecida, e, por isso, a IA, a depender do modelo adotado, atua mediante ontologias²¹⁵ e/ou análise de uma amostra dos dados²¹⁶ utilizada para o treino dos modelos²¹⁷, uma vez que a compreensão ou a representação da totalidade dos dados demandaria um poder mental (ou computacional) que é muitíssimo improvável de se alcançar.

Qualquer agente que queira computar o Direito como um todo está numa situação difícil, pois essa onisciência é um poder inalcançável, até mesmo para o deus *ex machina* da Inteligência Artificial. Reconhecer a pequenez do ser humano na condição de intérprete da realidade ajuda na compreensão de que sistemas de Inteligência Artificial judicial, embora limitados, podem auxiliar na compreensão do Direito. Assim, o homem não pode exigir poderes divinos de uma máquina, pois o instrumento é um auxiliar à humanidade, e nunca um substituto.

Ademais, tendo em vista a existência de prévia pesquisa com outra compreensão sobre a entropia jurídica, destaca-se que a visão da presente pesquisa se diferencia da apresentada por Rodrigo Greco²¹⁸, na medida em que a referência à

²¹³ “Pensamos situar-nos hoje num ponto crucial dessa aventura, no ponto de partida de uma nova racionalidade que não mais identifica ciência e certeza, probabilidade e ignorância” (PRIGOGINE, Ilya. Op. cit., p. 14).

²¹⁴ MIAILLE, Michel. **Introdução Crítica ao Direito**. 3ª ed. Tradução: Ana Prata. Lisboa: Editorial Estampa, 2005, p. 50-52.

²¹⁵ “An ontology is an ‘explicit, formal, and general specification of a conceptualization of the properties of and relations between objects in a given domain’ (Wyner, 2008). In other words, ontologies make concepts in a domain explicit so that a program can reason with them” (ASHLEY, Kevin D. **Artificial Intelligence and Legal Analytics**, p. 172).

²¹⁶ “Conceito fundamental: se olhar muito para um conjunto de dados, encontrará alguma coisa, mas isto pode não ser generalizável além dos dados para os quais está olhando” (PROVOST, Foster; FAWCETT, Tom. **Data Science para Negócios**. Rio de Janeiro: Alta Books, 2016, p. 15).

²¹⁷ LAGE, Fernanda de Carvalho. **Manual de Inteligência Artificial no Direito Brasileiro**. Salvador: Editora Juspodivm, 2021, p. 301-302.

²¹⁸ “Nesse sentido, entendemos que a tendência natural de aumento da entropia interna do sistema jurídico está relacionada à liberdade que os órgãos dotados de competência normativa possuem, segundo as normas secundárias, de estabelecer distintas conexões entre as normas que compõem o sistema jurídico” (GRECO, Rodrigo Azevedo. **Direito e Entropia**. Fortaleza: do autor, 2015, posição 941. E-book Kindle).

entropia feita pelo autor se refere ao papel das normas secundárias, numa leitura do pensamento de Herbert Hart, na maior ou menor organização do ordenamento jurídico no nível da produção normativa.²¹⁹ Isso que impacta no grau de liberdade do julgador e, segundo o autor, causaria uma insegurança jurídica²²⁰ a ser mitigada pela diminuição da entropia através da atuação do Supremo Tribunal Federal. A visão do referido autor aponta que, a depender do tipo de estruturação do ordenamento jurídico (piramidal, circular, em rede²²¹ ou caótica), haverá quase nenhuma ou grande entropia jurídica.²²²

Além disso, sobre a entropia no Direito, merecem referência os estudos dos seguintes pesquisadores: a) Roland Friedrich, Mauro Luzzatto e Elliott Ash²²³; b) Francesco Parisi²²⁴; c) Daniel Katz e Bommarito II²²⁵; d) Antony D'Amato²²⁶; e) Massimiliano Ferrara e Roberto Gaglioti²²⁷; f) Lyudmila Berg²²⁸; g) Craig Edgar²²⁹; h)

²¹⁹ “O subsistema de normas secundárias faz com que o sistema jurídico adquira uma estrutura qualquer, piramidal ou não, na medida em que limita, em maior ou menor grau, a liberdade do legislador e do juiz de se moverem em todas as direções do ordenamento, produzindo e relacionando as normas do sistema (gerais ou individuais) como bem entenderem” (Ibid., p. 950).

²²⁰ “Como vimos alhures, a redução da insegurança jurídica depende da diminuição da complexidade do ordenamento jurídico e, portanto, ou da diminuição da quantidade de normas jurídicas, ou da substituição das conexões circulares entre elas por conexões lineares. Essas medidas têm por objetivo fazer com que a estrutura do sistema caminhe paulatinamente em direção ao modelo piramidal, onde a segurança jurídica, sob o ponto de vista estrutural, é máxima e a entropia é mínima” (Ibid., posição 1375).

²²¹ A visão da estrutura em rede também remete ao problema das cadeias de validade do sistema jurídico, a partir dos pensamentos de Hans Kelsen e John Austin (RAZ, Joseph. **O Conceito de Sistema Jurídico**, p. 130-134).

²²² GRECO, Rodrigo Azevedo. Op. cit., posição 931.

²²³ FRIEDRICH, Roland; LUZZATTO, Mauro; ASH, Elliott. Entropy in Legal Language. **CEUR Workshop Proceedings**, Germany, v. 2645, p. 25-30, ago. 2020. DOI: 10.3929/ethz-b-000427800.

²²⁴ PARISI, Francesco. Entropy in Property. **The American Journal of Comparative Law**, v. 50, n. 3, p. 595-632, Summer 2002. DOI: 10.2307/841063.

²²⁵ KATZ, Daniel Martin; BOMMARITO, MJ. Measuring the complexity of the law: the United States Code. **Artificial Intelligence and Law**, Dordrecht, v. 22, p. 337-374, 16 sep. 2014. DOI: 10.1007/s10506-014-9160-8.

²²⁶ D'AMATO, Antony. Legal Uncertainty. **California Law Review**, Berkeley, v. 71, n. 1, p. 1-55, January, 1983. DOI: 10.2307/3480139.

²²⁷ FERRARA, Massimiliano; GAGLIOTI, Roberto Angelo. Legal Values and Legal Entropy: A Suggested Mathematical Model. **International Journal Of Mathematical Models And Methods In Applied Sciences**, v. 6, n. 3, p. 490-498, 2012.

²²⁸ BERG, Lyudmila N. System of Legal Impact from Social Entropy to Legal Development. **Revista Quid**, v. 2, Special Issue, p. 133-137, 2018.

²²⁹ EDGAR, Craig. Struggling Against Entropy: Monetary Patent Infringement Damages after eBay. **Drake Law Review**, Des Moines, v. 66, p. 45-67, 2018.

Jeanne Fromer²³⁰; i) Jopesh King²³¹; j) Jay Moran²³²; e k) Paul Stephan.²³³ Tais estudos têm o mérito de demonstrar o caráter entrópico do Direito, cada um ao seu modo e em sua área de enfoque, todavia não conduzem a uma visão sistemática de uma teoria da entropia jurídica.

Por outro lado, o autor Ted Sichelman²³⁴, a partir da teoria da informação de Claude Shannon²³⁵, buscou estabelecer um modelo matemático para tentar quantificar a entropia em sistemas jurídicos. O referido autor classifica a entropia dos sistemas jurídicos em entropia delimitativa (relacionada à criação da legislação, de regulamentos e de precedentes pelos órgãos competentes), entropia interpretativa (a ambiguidade de interpretação dos símbolos escritos em documentos jurídicos) e entropia aplicativa (a incerteza da aplicação de uma lei a um determinado conjunto fático).²³⁶ Desse modo, o autor busca um modelo matemático padrão para quantificar

²³⁰ FROMER, Jeanne C. An Information Theory of Copyright Law. **Emory Law Journal**, v. 64, n. 1, p. 71-128, 2014.

²³¹ KING JR, Joseph H. Deus ex Machina and the Unfulfilled Promise of New York Times v. Sullivan: Applying the Times for All Seasons. **Kentucky Law Journal**, Limestone Lexington, v. 95, n. 3, p. 650-674, 2007.

²³² MORAN, Jay P. Postmodernism's Misguided Place in Legal Scholarship: Chaos Theory, Deconstruction, and Some Insights from Thomas Pynchon's Fiction. **Southern California Interdisciplinary Law Journal**, Los Angeles, p. 155-199, 1997.

²³³ STEPHAN, Paul B. One Voice in Foreign Relations and Federal Common Law. **Virginia Journal of International Law**, v. 60, p. 1-79, 2019.

²³⁴ "What information costs reduce entropy? According to Shannon, the entropy of a system is precisely the number of informational bits needed to encode the microstates of the system. When the microstates themselves are unknown, such encoding involves an information cost in determining the precise informational bits of the microstates" (SICHELMAN, Ted. Quantifying Legal Entropy. **Frontiers in Physics**, Lausanne, v. 9, article 665054, 21 jun. 2021. DOI: 10.3389/fphy.2021.665054).

²³⁵ "The greater the uncertainty in microscopic configuration, the greater the entropy. In other words (and using log base 2 to capture the number of bits of entropy), $S = k \log_2 W$. Setting $k = 1$ for simplicity, and reducing W in the same manner as the Gibbs entropy, one arrives at the formula for Shannon [14] entropy in bits:

$$S = \sum p_i \log_2 p_i \quad (2)$$

If we suppose that each of the four microstates in our example are equally likely, the Shannon entropy of 10 cents is $-4 \times 0.25 \times \log_2 (0.25) = 2$ bits of information.³ This is sensible since there are only four choices, which we can label 00, 01, 10, or 11. This more generalized notion of entropy as the uncertainty over a range of informational microstates, as reflected in the Shannon entropy, will play a central role in quantifying legal entropy herein" (SICHELMAN, Ted. Op. cit.).

²³⁶ "Thus, it becomes incumbent to conceptualize the different domains of entropy that arise in legal systems. Delineative entropy involves the ambiguity and related information costs in formulating the law in the first instance, typically into written symbols, in a constitution, statutes, regulations, judicial decisions, and the like. As noted earlier, interpretive entropy concerns the ambiguity in interpreting the written symbols in legal documents, including not only constitutions, statutes, and regulations, but also judicial decisions. Such an endeavor is not merely textual in nature, but will often involve relying upon institutional and social norms, which themselves can be uncertain. Finally, applicative entropy is roughly the uncertainty involved in applying an interpreted law to a given set of facts" (Id.).

a entropia interpretativa²³⁷ e a entropia aplicativa²³⁸, o que é digno de reconhecimento na medida em que pode ser utilizado como um método de apoio aos sistemas de IA jurídicos.²³⁹ Todavia, Sichelman é sincero ao concluir que “o modelo não aborda detalhes práticos importantes sobre como avaliar as probabilidades subjacentes necessárias para calcular a entropia jurídica” [tradução livre].²⁴⁰

Noutro ângulo, para evitar a frustração de tentar medir o Direito com um único modelo matemático, bem como tendo por base a busca de respostas através da Teoria do Direito, a premissa aqui desenvolvida leva em consideração que a entropia é o estado do Direito em si mesmo (pois o jurídico é parte de um todo, do próprio Universo). Não somente um ordenamento jurídico, um sistema normativo ou um bloco de enunciados dogmáticos ou proposições jurídicas, mas também um conjunto de significados (e, por trás destes, dados e informação) contextualizados através de sua argumentação (pautada na razão prática ou argumentação jurídica como um caso especial do discurso prático geral²⁴¹). Dessa forma, a concepção de entropia jurídica trazida aqui atua num nível mais profundo, tratando da complexidade da existência do

²³⁷ “If the only step involved in this process were to interpret the express text of a legal rule with a standard dictionary, then techniques using word and related forms of linguistic entropy would provide a fairly accurate value of the interpretive entropy. For instance, one could measure the word entropy of legal texts that measures the ambiguity inherent in each word using a standard corpus (cf. Piantadosi et al. [54]).

Yet, legal interpretation extends well-beyond textual interpretation with a standard dictionary. As an initial matter, many legal terms are ‘terms of art’, requiring interpretation by specialized, legal dictionaries. Quantifying interpretive entropy using text-based measures of entropy must rely therefore not on a standard corpus, but a specialized one. More problematic, legal interpretation typically draws upon the language of other legal rules, which necessitates determining how those other rules’ affect the entropy of the rule-at-issue. And even more problematic, interpreting rules draw on yet more disconnected sources, such as legislative history, judicial decisions, and even general policies and social norms. Not to mention that interpretive entropy applies not only to the interpretation of legal rules, but also legal documents more generally, such as patents and contracts, which introduce further interpretive issues” (Id.).

²³⁸ “Once a legal rule has been interpreted, in order to understand how it specifically regulates human behavior, it must be applied to a specific situation, or set of facts.¹⁶ Even though the applicable legal rule has been fully interpreted, residual indeterminacy in the application of the law may remain and can be quantified by applicative entropy. This indeterminacy can arise from ‘uncertainty as to the impact evidence will have on the decisionmaker,’ idiosyncratic behavior in adjudication by a decisionmaker such as a judge or jury, and the influence of extra-legal factors on the regulatory and judicial process [56].

Applying the law typically results in the imposition of liability (or no liability), plus some form of remedy in the event liability is imposed. Again, Shannon entropy can be used to measure the entropy of liability and the conditional entropy to measure the entropy present in the range of remedies in the event liability is imposed. Before turning to these specifics, it is instructive to examine the typology of Hohfeld [15], as it offers a sound, quantitative conceptual basis to describe the entropy of composite legal systems and subsystems, which is illustrated well by applicative entropy” (Id.).

²³⁹ “First, it offers a potential template for how legal AI systems can measure and store information about the uncertainty of legal systems” (Id.).

²⁴⁰ Id.

²⁴¹ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 188-189.

Direito e da complexidade (e insondabilidade) total do fenômeno jurídico, ainda que um hipotético ordenamento tenha, eventualmente, sido estaticamente organizado num plano hierárquico para limitar ou impedir flutuações interpretativas, ou ainda que (um dia, no futuro) o Direito possa ser medido num dado recorte do tempo por uma determinada técnica.

A certeza do Direito é um recorte temporal idealizado²⁴², uma vez que a entropia jurídica promove a incerteza²⁴³ na interpretação jurídica. Desse modo, as supostas certezas decorrentes de processos explicáveis previamente²⁴⁴ são uma distribuição difusa²⁴⁵ de possibilidades²⁴⁶ interpretativas num futuro entrópico²⁴⁷ (atrelado à natural imperfeição do discurso real²⁴⁸) que destrói e reconstrói²⁴⁹ as relações jurídicas. Nesse sentido, a irreversibilidade da entropia jurídica também tem um papel

²⁴² “Essa física tradicional unia conhecimento completo e certeza: desde que fossem dadas condições iniciais apropriadas, elas garantiam a previsibilidade do futuro e a possibilidade de retrodizer o passado. Desde que a instabilidade é incorporada, a significação das leis da natureza ganha um novo sentido. Doravante, elas exprimem possibilidades” (PRIGOGINE, Ilya. Op. cit. p. 12).

²⁴³ “O conceito de correção sublima-se, por um lado, a uma ideia regulativa, por outro, ele é relativizado em alta medida e unido com incerteza. Esse resultado, contudo, não é a última palavra, se a tese é acertada, que o valor prático da teoria do discurso mostra-se em sua totalidade, primeiro então, quando ela é transformada em uma teoria de base das instituições, apoiadas na discussão, do estado constitucional democrático. Fundamentar essa tese iria significar, contudo, a abertura de um novo tema” (ALEXY, Robert. **Direito, Razão, Discurso**, p. 100).

²⁴⁴ “Os processos reversíveis, em compensação, correspondem sempre a idealizações: devemos negligenciar a fricção para atribuir ao pêndulo um comportamento reversível, e isto só vale como uma aproximação” (PRIGOGINE, Ilya. Op. cit., p. 25).

²⁴⁵ “A cada transição, o futuro é, portanto, incerto. Contudo, no nível estatístico, o modelo fornece um comportamento regular bem definido, correspondente a uma difusão” (Ibid., p. 47).

²⁴⁶ “Como já ressaltamos, tanto na dinâmica clássica quanto na física quântica, as leis fundamentais exprimem agora possibilidades e não mais certezas. Temos não só leis, mas também eventos que não são dedutíveis das leis, mas atualizam as suas possibilidades” (Ibid., p. 13).

²⁴⁷ [Pois] “os fenômenos de difusão são fenômenos irreversíveis: a difusão corresponde a uma aproximação da uniformidade no futuro e produz entropia” (Ibid., p. 45).

²⁴⁸ “A quarta relatividade, a relatividade com vistas a *datas*, é inevitável perante os limites de cada discurso real. Em resultados só discursivamente possíveis, portanto, dissensos, a ideia regulativa da correção exige o caráter não definitivo do resultado. Em resultados discursivamente necessários, portanto, consensos, o caráter necessariamente provisório resulta da imperfeição necessária de cada discurso real” (ALEXY, Robert. **Direito, Razão, Discurso**, p. 99).

²⁴⁹ “Em primeiro lugar, nossa recusa da banalização da irreversibilidade funda-se no fato de que, mesmo na física, a irreversibilidade não pode mais ser associada apenas a um aumento da desordem. Muito pelo contrário, os desenvolvimentos recentes da física e da química de não equilíbrio, mostram que a flecha do tempo pode ser uma fonte de ordem” (PRIGOGINE, Ilya. Op. cit., p. 29).

construtivo²⁵⁰, haja vista que o intérprete muda o mundo jurídico ao observá-lo²⁵¹, impulsionando a própria entropia do Direito.

Sem a consideração da entropia, escolhas doutrinárias tenderão a responder às questões com a simples aderência a uma ou outra teoria. Uma determinada decisão será considerada positivista, outra estará ligada ao modelo de integridade de Dworkin, uma outra, ainda, será relacionada ao “neconstitucionalista”, e, por fim, outra será “garantista”. Ao final, a própria interação entre as diversas teorias e seus reflexos, maiores ou menores, na forma como cada Juiz decide, causam entropia jurídica. Assim, como a diferença entre a bola na rede e a bola na trave pode ser de um centímetro num jogo de futebol (e tudo à volta influencia nesse centímetro), a diferença entre a procedência ou a improcedência de um pedido no processo, em casos limites, pode ser uma palavra, uma imagem ou um pensamento a mais (ou a menos) na mente do julgador.

O ponto aqui discutido consiste em como os esforços de coesão voltados à estabilização racional do Direito são importantes e devem continuar existindo, mas não conseguem sequer perceber a entropia jurídica que atua indiscriminadamente contra o esforço de previsibilidade²⁵² que o Direito busca.

Como exemplo disso, o leitor pode agora pensar num personagem, o nome dele é José. Após duras penas e muito trabalho José compra uma casa para viver com a sua família. A vida é dinâmica e José tenta ganhar algum dinheiro empreendendo, motivo pelo qual abre uma pequena empresa. Contudo, no contexto

²⁵⁰ “Longe do equilíbrio, o papel construtivo da irreversibilidade torna-se ainda mais impressionante. Ela cria, ali, novas formas de coerência”. (...) “A vida só é possível num universo longe do equilíbrio”. (...) “1 Os processos irreversíveis (associados à flecha do tempo) são tão reais quanto os processos reversíveis descritos pelas leis tradicionais da física; não podem ser interpretados como aproximações das leis fundamentais.

2 Os processos irreversíveis desempenham um papel construtivo na natureza.

3 A irreversibilidade exige uma extensão da dinâmica” (Ibid., p. 30-31).

²⁵¹ “Outras questões estão diretamente ligadas ao problema do tempo. Uma é o papel estranho conferido ao observador na teoria quântica. O paradoxo faz de nós os responsáveis pela quebra de simetria temporal observada na natureza”. (...) “Na interpretação usual a medição, que impõe uma referência ao observador na teoria quântica, corresponde a uma quebra de simetria temporal. Em compensação, a introdução de instabilidade na teoria quântica leva a uma quebra na simetria do tempo” (Ibid., p. 12-13).

²⁵² “Se o nosso mundo devesse ser compreendido por meio do modelo de sistemas dinâmicos estáveis, não teria nada em comum com o mundo que nos cerca: seria um mundo estático e predizível, mas não estaríamos lá para formular as predições. No mundo que é o nosso, descobrimos em todos os níveis flutuações, bifurcações, instabilidades. Os sistemas estáveis que levam a certezas correspondem a idealizações, a aproximações” (Ibid., p. 59).

da crise econômica de 2009, a empresa de José vai à falência²⁵³, como tantas outras empresas. Após isso, José passa a responder por dívidas da empresa, uma vez que foi o garantidor de diversos contratos e negócios realizados pela pessoa jurídica. Em 2011, o Banco X ajuizou uma ação de execução de uma das dívidas de José, sendo que ao tentar penhorar aquela única casa do devedor (que fica numa cidade pequena do interior), descobriu que José passou a morar, por anos, na casa de um filho na capital do Estado, pois precisava de acesso ao médico para tratar um câncer que lhe acometeu. A cidade da residência do filho de José tem o melhor tratamento contra a doença, sendo justificada sua mudança.

O Banco X, ao ter acesso à declaração de imposto de renda do devedor, presente nos autos do processo devido a uma ordem judicial, justificou a penhora do único imóvel de José no fato de que o mesmo teria sido “abandonado” e não gerava qualquer renda para José, apoiando-se necessidade de “moradia permanente” prevista no art. 5º da Lei 8.009/1990 e nos precedentes do Superior Tribunal de Justiça baseados na Súmula 486 do mesmo Tribunal, a qual prevê que é “impenhorável o único imóvel residencial do devedor que esteja locado a terceiros, desde que a renda obtida com a locação seja revertida para a subsistência ou a moradia da sua família”.²⁵⁴ José se defendeu da alegação do Banco X dizendo que nunca abandonou o imóvel e que, quando curado de seu câncer, quer retornar para sua vida pacata do interior, motivo pelo qual, pelas esperanças de cura, nunca alugou o imóvel para dele obter renda imobiliária. Do lado do Banco X há uma situação fática e jurídica consolidada na busca de proteção de seu direito de credor, já ao lado de José se apresentam peculiaridades nunca antes tratadas pelos precedentes, mas que poderiam ser justificadas em fundamentos constitucionais, como os direitos à dignidade humana, à vida e à saúde. Drama jurídico e entropia jurídica andam juntos, por isso a IA judicial deve atuar para ampliar a visão do julgador aos inúmeros argumentos que brotam do caso, nunca como simples automatização que não irá reverberar a entropia das soluções possíveis.

O que sobressai desse exemplo é a variabilidade inerente à entropia jurídica, que resulta em possibilidades interpretativas que inserem um grau de

²⁵³ Aqui o termo “falência” é utilizado num contexto mais amplo, para além do estritamente conceito jurídico.

²⁵⁴ BRASIL. Superior Tribunal de Justiça. **Súmula 486 do STJ**. Disponível em: <https://www.stj.jus.br/publicacaoinstitutional/index.php/sumstj/article/download/5174/5299>. Acesso em: 20 jul. 2022.

imprevisibilidade e que podem ser delimitadas, para fins auxílio à compreensão, mediante a classificação da entropia jurídica em entropia fática, entropia político-social e entropia interpretativa.

A **entropia fática** se apresenta no mundo sensível, é explicada pela física e pelas demais ciências exatas, o que traz a variabilidade do mundo fático²⁵⁵ como um desafio à aplicação das normas jurídicas e à consistência dos sistemas jurídicos. Os argumentos práticos do tipo geral são diretamente influenciados pela entropia fática. Por sua vez, a **entropia político-social** é a causada pela ação do ser humano, na condição de indivíduo que atua numa sociedade não atomizada, na qual o agir humano e o contexto social repercutem na variabilidade e na imprevisibilidade da interpretação do Direito. Já a **entropia interpretativa** consiste na variabilidade de formas de interpretação do Direito, em decorrência dos posicionamentos de cada julgador individualmente considerado (com uso de doutrina, precedentes, argumentos e teorias do Direito), bem como pela interação entre os diversos posicionamentos individuais nos órgãos de julgamento colegiado (Câmaras, Turmas e Tribunais).

Essas três espécies de entropia compõem um tridente, que forma a entropia jurídica e rasga o tecido do ordenamento, dos sistemas jurídicos, das normas e suas interpretações. Construção, desconstrução e reconstrução, num movimento típico do Direito e de sua formação, o que repercute na sua manifestação interpretativa.

Dessa maneira, graças à entropia que flui no que se convencionou chamar de dilúvio de dados, toda informação é um recorte contextualizado, e é a renúncia aos dados não conhecidos. Sendo assim, medir os dados do mundo e contextualizá-los traz informação enquadrada (como uma pintura, uma *Monalisa* de Leonardo Da Vinci). Informação bem interpretada leva à sabedoria. Logo, num enquadramento na teoria da informação, a aplicação do Direito demanda sabedoria. Contudo, a sabedoria nasce engarrafada na informação solidificada com os dados disponíveis ou utilizados para a representação. Assim, o Direito, fiel ao mundo e à época em que se manifesta²⁵⁶, é único e pode ser apreciado em sua singularidade (quase como se fosse um suave vinho tinto).

²⁵⁵ “No nível estatístico, as ressonâncias acarretam a ruptura do determinismo, introduzem a incerteza no contexto da mecânica clássica e quebram a simetria do tempo” (PRIGOGINE, Ilya. Op. cit., p. 47).

²⁵⁶ “*Hic Rhodus, hic saltus*. A missão da filosofia está em conceber o que é, porque o que é a razão. No que se refere aos indivíduos, cada um é filho do seu tempo; assim também para a filosofia que, no pensamento, pensa o seu tempo. Tão grande loucura é imaginar que uma filosofia ultrapassará para fora do seu tempo, transporá *Rhodus*. Se uma teoria ultrapassar estes limites, se construir um mundo tal como entenda dever ser, este mundo existe decerto, mas apenas na opinião, que é um elemento

Será que a obra *Monalisa* poderia ser melhorada? Ou o melhor vinho ainda não foi conhecido? Ver todos os quadros do mundo é não ver os quadros que ainda não foram pintados. Ainda que fosse possível beber todos os vinhos do mundo²⁵⁷, os vinhos experimentados são o limite para não se saborear os vinhos que poderiam existir. Em suma, interpretar aquilo que é conhecido (informação) é uma renúncia àquilo que ainda não é conhecido. Desse modo, uma das funções da Inteligência Artificial no Direito é explorar novos caminhos, nuances e significados que possam passar despercebidos pelos sentidos humanos, de maneira a ampliar a banda de processamento por meio da colaboração inerente à relação humano-máquina. Assim, a entropia jurídica serve como uma advertência ao intérprete humano, pois, diante de suas limitações, o auxílio de Inteligência Artificial pode ser útil. Em suma, frente à entropia jurídica o intérprete humano deve ser humilde, conhecer suas limitações e potenciais e buscar o auxílio que a tecnologia lhe possibilita.

Interpretar é um esforço de boa-fé ao buscar ver aquilo que se apresenta, o que também pode ser auxiliado pela análise de dados.²⁵⁸ Todavia, quando se interpreta o mundo, se renuncia ao Universo que não se despe perante os olhos da humanidade. Nunca será bebido o maravilhoso vinho que se perdeu nas aleatoriedades da entropia.²⁵⁹ Por isso, toda interpretação é uma simplificação. Dessa forma, o ser humano não é a medida de todas as coisas, o ser humano apenas reduz todas as coisas à sua capacidade de processamento. Portanto, o intérprete pode desfragmentar tudo em dados e informações. Mas deve saber que se trata de uma parcela da realidade, pois os fragmentos sozinhos não fazem sentido, já que um fragmento de um objeto não forma o mesmo.

Nesse sentido, se a busca por uma suposta “verdade”²⁶⁰ hermenêutica terminar em um certo recorte da realidade, o intérprete será como um náufrago agarrado às próprias simplificações do mundo como restos de sua embarcação, os quais a

inconsciente sempre pronto a adaptar-se a qualquer forma” (HEGEL, Georg Wilhelm Friedrich. **Princípios da Filosofia do Direito**. Tradução: Orlando Vitorino. São Paulo: Martins Fontes, 1997, p. XXXVII).

²⁵⁷ Isso sim seria o poder de um Deus.

²⁵⁸ “Conceito fundamental: *a partir de uma grande massa de dados, a tecnologia da informação pode ser usada para encontrar atributos descritivos informativos de entidades de interesse*” (PROVOST, Foster; FAWCETT, Tom. Op. cit, p. 14).

²⁵⁹ Veja-se que o exemplo é meramente didático. Este autor gosta da analogia com vinhos, mas outros exemplos podem ficar por conta imaginação do leitor.

²⁶⁰ A partir do conceito de entropia jurídica desenvolvido nesta obra, desde já deve ser referido que **não se defende aqui “uma” ou “a” “verdade”**, sendo que a referência se dá no contexto da investigação jurídico-científica das possibilidades de interpretação.

hermenêutica chama de pré-conceitos.²⁶¹ Sendo assim, o círculo hermenêutico²⁶² é o eterno processamento do intérprete na busca de esculpir a madeira da sua embarcação (o conceito) com base naquilo que a realidade lhe impõe como mundo da vida. Contudo, os dados mudam e a “tempestade” pode vir mais forte da próxima vez. E quando o conceito falha (não mais dá conta da realidade) o que o hermeneuta faz? Volta ao círculo hermenêutico, olha para o passado (para os pré-conceitos e para o mundo), assim, nasce um conceito novo feito daquela “madeira antiga”.²⁶³

Essa é a aventura do hermeneuta, projetar conceitos olhando para o passado e desafiar o oceano revolto do futuro bradando que agora chegou mais próximo da “verdade”.²⁶⁴ Ondas virão! O “barco” se quebra e lá vai o hermeneuta invocar o círculo hermenêutico. Portanto, é necessária humildade ao “construtor de barcos”, pois o “oceano” é cheio de armadilhas. Quando o barco se partir não bastará bradar que os problemas são as “ondas” muito revoltas do mundo da vida ou que a realidade já não deixa a sua cambaleante interpretação parar em pé.

²⁶¹ “Heidegger, pelo contrário, descreve esse círculo de uma forma tal que a compreensão do texto se encontre determinada, continuamente, pelo movimento de concepção prévia da pré-compreensão. O círculo do todo e das partes não se anula na compreensão total, mas nela alcança sua mais autêntica realização” (GADAMER, Hans-Georg. **Verdade e Método**: Traços fundamentais de uma hermenêutica filosófica. 3ª ed. Tradução: Flávio Paulo Meurer. Petrópolis: Editora Vozes, 1999, p. 439).

²⁶² “Como já afirmado, o círculo hermenêutico tem origem no método fenomenológico de HEIDEGGER e pressupõe a ideia de mundo como horizonte de sentido onde se mora, lugar de manifestação do possível, da linguagem.

Ele aparece aliado ao grupo de categorial de existenciais heideggerianos, tais como a consciência de finitude, o diálogo que permite a compreensão e se realiza pela abertura (possibilidade de compreender) do ente que, por sua vez, depende da disposição, seu modo de ser enquanto projeto, da temporalidade (que se revela em função da consciência de finitude) e historicidade (instância fundante da história colocada pela temporalidade como marca ineliminável da existência humana).” (...)

“Portanto, o círculo revela um movimento ontológico (ao compreender o ente homem, se compreende e ‘ser’ e vice-versa, simultaneamente), que se realize na fórmula sujeito-abertura-estranhamento/internalização ôntica-sujeito, numa circularidade que toma forma espiral porque não deve se fechar sobre si mesmo” (MARRAFON, Marco Aurélio. **O Caráter Complexo da Decisão em Matéria Constitucional**: Discursos sobre a Verdade, Radicalização Hermenêutica e Fundação Ética na Práxis Jurisdicional. Rio de Janeiro: Lumen Juris, 2010, p. 119-120).

²⁶³ O caso difere do paradoxo do Navio de Teseu, proposto por Plutarco, na medida em que naquela alegoria a embarcação é reparada com madeira nova e surge o questionamento se o navio ainda é o mesmo. Já no caso aqui tratado se tem se refere ao uso de pré-conceitos (madeira antiga) para dar respostas aos problemas novos sem uma adequada compreensão da complexidade entrópica da realidade.

²⁶⁴ E este autor repete, não se debate “a verdade” nesta obra, sendo a referência à hermenêutica filosófica apenas um exemplo didático de como a entropia jurídica desvela outras possibilidades ao intérprete do Direito, seja este um eterno buscador da “verdade” hermenêutica ou um relativista do Direito (onde inúmeras escolas de pensamento poderiam ser referidas).

Mas onde reside a limitação epistêmica de pensamentos como esse? A resposta está na não inclusão da entropia²⁶⁵ em sua abordagem.

Assim, o Direito é um processo comunicacional complexo²⁶⁶, complexidade em que a parte implícita²⁶⁷ de sua aplicação, inerente à subjetividade da mente do intérprete²⁶⁸, é inalcançável.

Uma demarcação inicial deve ser realizada em relação à principal linha de pensamento filosófico que tem se posicionado pelo abandono da ideia²⁶⁹ de inclusão da Inteligência Artificial no processo de decisão judicial.²⁷⁰ Em comparação com o modelo de cooperação humano-máquina²⁷¹ e pretensão de correção aqui pensado, ao não levar em consideração a entropia do mundo na busca pelas respostas corretas, a hermenêutica filosófica tem uma noção de tempo transcendental, atribuindo à compreensão da temporalidade do “agora” a alcunha de “vulgar”, conforme referido por Martin Heidegger²⁷². Dentre os estudos clássicos do pensamento hermenêutico,

²⁶⁵ “A irreversibilidade não pode mais ser identificada com uma mera aparência que desapareceria se tivéssemos acesso a um conhecimento perfeito. Ela é uma condição essencial de comportamentos coerentes em populações de bilhões de bilhões de moléculas” (PRIGOGINE, Ilya. Op. cit., p. 11).

²⁶⁶ “A argumentação diante do tribunal (da mesma forma que outros tipos de discussão jurídica, tais como aconselhamentos judiciais, discussões dogmáticas, comentários da lei etc.) difere de discursos práticos gerais, dado seu vínculo com o direito vigente, e dadas também as restrições especiais de ordem processual que faça jus a exigência de uma decisão autorizada e do direcionamento ao êxito, por parte de cada uma das partes litigantes. Não obstante, a argumentação em juízo contém elementos essenciais que só podem ser apreendidos de acordo com o modelo de argumentação moral ou com o modelo da discussão sobre a correção das declarações normativas. Por isso – estejam elas relacionadas ou a questões do direito e da moral ou hipóteses científicas e obras de arte –, todas as argumentações exigem a mesma forma básica de organização relativa à procura cooperativa da verdade, que subordina os meios da -erística ao objetivo de gerar convicções intersubjetivas, em virtude dos melhores argumentos” (HABERMAS, Jürgen. **Teoria do Agir Comunicativo**, 1, p. 78-79).

²⁶⁷ “Assistimos ao surgimento de uma ciência que não mais se limita a situações simplificadas, idealizadas, mas nos põe diante da complexidade do mundo real, uma ciência que permite que se viva a criatividade humana como a expressão singular de um traço fundamental comum a todos os níveis da natureza” (PRIGOGINE, Ilya. Op. cit., p. 15).

²⁶⁸ “A human being that is undertaking a legal reasoning task has some form of implicit reasoning that is hidden within their mind” (ELIOT, Lance B. **AI and Legal Reasoning Essentials**. Stanford: LBE Press, 2020, p. 43)

²⁶⁹ “Na verdade, chamar esses grupos de ‘ludditas fundamentalistas’ é redundante, porque o ludditismo é inerentemente fundamentalista. Ele reflete a ideia de que a humanidade ficará melhor sem mudanças, sem progresso” (KURZWEIL, Ray. **A Singularidade Está Próxima**, p. 473).

²⁷⁰ STRECK, Lenio; BERNST, Luísa Giuliani; GOMES, Jefferson de Carvalho. Inteligência Artificial: mesmos problemas, mas na versão Hi-Tech. **Constituição, Economia e Desenvolvimento: Revista da Academia Brasileira de Direito Constitucional**, Curitiba, 2021, vol. 13, n. 25, p. 333-342, ago./dez., 2021.

²⁷¹ “O desempenho desses sistemas cada vez mais se baseia na integração de múltiplos tipos de inteligência artificial (IA). Mas, no momento em que a IA deixa a desejar em alguma dessas áreas de trabalho, céticos apontam essa área como um inerente bastião da permanente superioridade humana em relação à capacidade de nossas próprias criações” (KURZWEIL, Ray. Op. cit., p. 24).

²⁷² “A temporalidade ekstática e horizontal temporaliza-se, *primordialmente*, a partir do *porvir*. A compreensão vulgar do tempo, ao contrário, vê o fenômeno fundamental no *agora* e no puro agora que, moldado em toda sua estrutura, se costuma chamar de ‘presente’. Daí se pode depreender que, em

embora a temporalidade seja um de seus objetos²⁷³ não se verifica uma preocupação com o processo estocástico de modificação da facticidade e seu análogo efeito sobre o Direito.

A partir dessa abordagem, a historicidade, dentro do contexto da temporalidade delineado por Gadamer²⁷⁴, é uma amarra da interpretação ao momento de mundo que se apresenta, à sua tradição e aos limites da pré-compreensão, mas não leva em conta o processo de explicação que surge a partir do reconhecimento da entropia naturalmente atrelada à flecha do tempo. Em suma, para a hermenêutica filosófica a temporalidade e a historicidade²⁷⁵ são atributos para se compreender aquilo que é, mas não necessariamente levam em conta os efeitos da entropia sobre o “ser”.²⁷⁶ Desse modo, a entropia é tacitamente deixada à margem pela hermenêutica filosófica, na medida em que tradicionalmente é estudada sob o foco das ciências exatas, regidas pelo tipo de método científico que Gadamer repele quando se refere às ciências humanas.²⁷⁷ Essa separação forte entre as metodologias das ciências

princípio, deve ficar fora de qualquer possibilidade esclarecer e, sobretudo, derivar *desse agora* o fenômeno ekstático e horizontal do *instante* que pertence à temporalidade própria. De modo correspondente, não se confundem o porvir ekstático, o ‘então’ datável da significância e o conceito vulgar de futuro, no sentido de simples agoras que ainda não advieram e que estão em advento. Tampouco coincidem o vigor de ter sido, ekstaticamente compreendido, o ‘outrora’ datável da significância e o conceito do passado, no sentido dos puros agora passados. O agora já não fica grávido do agora-ainda-não. Ao contrário, a atualidade surge do porvir na unidade ekstática e originária da temporalidade” (HEIDEGGER, Martin. **Ser e Tempo**. Tradução: Márcia Sá Cavalcante Schuback. 4ª ed. Petrópolis: Vozes, 2009, p. 524).

²⁷³ Especialmente para na busca de se explicar os conceitos de Dasein, “ser-aí-no-mundo”, presença, preconceitos e historicidade, típicos da hermenêutica filosófica.

²⁷⁴ “Aqui temos o ponto, no qual a pesquisa filosófica é remetida hoje a novos caminhos. Em seu livro *Ser e Tempo*, um livro que fez época, Martin Heidegger colocou a historicidade do ser-aí humano tão radicalmente no centro da filosofia que, juntamente com as ilusões da razão histórica (Dilthey), também se tornou caduca a supressão desesperada das ilusões por parte de Nietzsche. O ser do ser-aí humano é um ser histórico. Isso não significa que ele está presente como o ser-aí da natureza, apenas de maneira mais frágil e mutável do que ela. Tampouco significa que ele se sabe como histórico, que ele possui uma consciência historiológica. Ao contrário, ele só possui uma tal consciência porque é histórico. Ele é o seu futuro, a partir do qual ele se temporaliza em suas possibilidades. Todavia, o seu futuro não é o seu projeto livre, mas um projeto jogado. Aquilo que ele pode ser é aquilo que ele foi. O ideal de compreensão histórica universal é uma falsa abstração que esquece a historicidade” (GADAMER, Hans-Georg. **Hermenêutica em Retrospectiva**. Tradução: Marco Antônio Casanova. Petrópolis: Vozes, 2009, p. 191-192).

²⁷⁵ “A palavra ‘historicidade’ visa algo diverso daquilo que a ciência da história do universo tem por objeto, por mais abrangente que essa história seja pensada” (Ibid., p. 222).

²⁷⁶ “A primeira consequência dessa assimetria é que, é que, nas mentes das pessoas, a relação entre passado e futuro não aprende com a relação entre o passado e o passado que o antecedeu. Existe um ponto cego: quando pensamos no amanhã, não o enquadrados nos termos do que pensamos ontem ou anteontem. Por causa desse defeito introspectivo deixamos de aprender sobre a diferença entre nossas provisões passadas e os resultados subsequentes. Quando pensamos sobre o amanhã, simplesmente o projetamos como um outro ontem” (TALEB, Nassim N. **A Lógica do Cisne Negro: O Impacto do Altamente Improvável**. Tradução: Marcelo Schild. 1ª ed. Rio de Janeiro: BestBusiness, 2019, p. 248-249).

²⁷⁷ GADAMER, Hans-Georg. **Hermenêutica em Retrospectiva**. p. 233-237.

humanas e das ciências exatas²⁷⁸ é um dos motivos pelos quais a entropia do Direito vem sendo deixada de lado.

A historicidade é entrópica, pois também é regida por Cronos²⁷⁹, e assim como o titã, devora lentamente seus filhos. O processo pode ser rápido ou mais lento, segundo o ponto de vista humano, todavia ele é inexpugnável, imparável e implacável. Para além disso não existe um determinismo²⁸⁰ e uma determinabilidade, pois tudo isso é limitado pela aleatoriedade.

A enviesada compreensão da inclusão da Inteligência Artificial no Direito que tem sido sustentada por parte da corrente da hermenêutica filosófica (mediante desconfiança na tecnologia de apoio a decisão judicial²⁸¹ ou por meio de excessiva exigência de uma IA forte voltada à “hermenêutica artificial”²⁸²) tem uma razão de fundo. Esta se dá porque referidas críticas não se preocupam com o fato de que a realidade se reverte em informação e, em última análise, em dados. A entropia é a medida da ignorância sobre algo²⁸³, pois até mesmo os fechamentos interpretativos são estados iniciais que evoluem para a desordem da entropia jurídica.

²⁷⁸ “Aqui não se coloca de maneira alguma o problema do relativismo. O que se encontra à base desse problema é o conceito de verdade da ciência e a objetivação de todos os âmbitos de pesquisa. Isso, porém, não instaura nenhuma relação de companheirismo, e uma tal relação é para nós imposta pelas condições de nossa temporalidade e de nossa historicidade” (Ibid., p. 275).

²⁷⁹ “Cronos sintetiza a sucessão irreversível entre antes, agora e depois, na qual tudo aquilo que já ocorreu não pode ser desfeito e nada daquilo que está por vir pode ser conhecido plenamente.” (...) “Para os gregos, Cronos era a personificação do tempo linear, que destrói, conduz até a morte, um deus gigante que devorava seus próprios filhos (como no quadro *Saturno devorando a un hijo*, de Francisco de Goya) temendo perder seu poder. Cronos, que constitui a raiz do termo cronômetro é o maior símbolo da temporalidade sucessiva, computável. Uma temporalidade quantitativa, o tempo: “... da viagem que conduz do nascimento até a morte e marca, também, o início e o fim de cada lapso de nossa vida, independentemente se ditos fragmentos temporais foram plenos de tempo, ou este apenas foi perceptível em seu passar” (LORENZETTO, Bruno Meneses. **O Silêncio das Sereias: Tempo, Direito e Violência na Modernidade**. Dissertação (Mestrado em Direito). Universidade Federal do Paraná (UFPR), Curitiba, 2010, p. 12 e p. 21).

²⁸⁰ “Num mundo isomorfo e um conjunto de corpos sem interação, não há lugar para a flecha do tempo, nem para a auto-organização, nem para a vida” (PRIGOGINE, Ilya. Op. cit., p. 42).

²⁸¹ STRECK, Lenio; BERNST, Luísa Giuliani; GOMES, Jefferson de Carvalho. Op. cit.

²⁸² “A possibilidade, como o texto propõe, de uma hermenêutica artificial, portanto, talvez não seja tão futurista assim. Com mecanismos de *machine learning*, e com a evolução do *deep learning* (ou aprendizado de máquina profundo), essa possibilidade tende, progressivamente, a se tornar realidade. Somente com o alcance da I.A. geral ou forte (dotada intrinsecamente de um pensar hermenêutico) será possível a construção de decisões judiciais que, respeitando, por exemplo, coerência, integridade, princípios de direito etc., forneçam as respostas adequadas ao direito, as respostas mais corretas possíveis” (HOMMERDING, Adalberto Narciso; HARTMANN, Gabriel Henrique. Direito, Hermenêutica e Inteligência Artificial: Construindo Pontes entre Decisão Judicial, Compreensão Existencial e Mecanismos de Machine Learning. **Quaestio Iuris**, v. 14, n. 4, Rio de Janeiro, p. 1001-1038, 2021. DOI: 10.12957/rqi.2021.58084).

²⁸³ “En realidad, la definición técnica de la entropía (medida en la unidad más conveniente, la llamada constante de Boltzmann) está íntimamente relacionada con el número de maneras diferentes en que un sistema cerrado puede encontrarse en un macroestado particular. La entropía de un sistema en un macroestado dado es, aproximadamente, la cantidad de información — el número de bits — necesaria

Destaca-se que não se busca, aqui, desautorizar ou refutar a teoria hermenêutica como uma das possíveis teorias jurídicas dentro do aspecto de doutrina razoável.²⁸⁴ Apenas se aponta sua insuficiência para a compreensão das formas de utilização adequadas de sistemas Inteligência Artificial na aplicação do Direito.

Assim, a maneira mais segura de se realizar uma aferição da correção da aplicação do Direito é através dos argumentos expostos pelo julgador, pois, assim, se garante um mínimo grau de objetividade. Isso ocorre porque a legitimidade do agir comunicativo não está no todo imaculado e inalcançável de uma mensagem, mas na possibilidade de, com os signos necessários²⁸⁵, a mensagem ser codificada, transmitida, decodificada e recebida.²⁸⁶

Dessa forma, uma interpretação a respeito de algo nunca é feita pela total compreensão de cada mínimo detalhe, pois são utilizadas abstrações a respeito do que é interpretado.²⁸⁷

Sendo assim, a interpretação é, em certa medida, um ato de comunicação entre o legislador e o intérprete do Direito²⁸⁸, bem como entre esse e o destinatário da interpretação (o sujeito de Direito). Toda vez que o legislador (ou o constituinte) cria uma norma jurídica que busca a generalidade e a abstração²⁸⁹ há a previsão de um

para especificar uno de los microestados correspondientes al macroestado, considerando todos los microestados como igualmente probables" (GELL-MAN, Murray. **El Quark y el Jaguar**: Aventuras en lo simple y lo complejo. Traducción: Ambrosio García Leal y Romualdo Pastor. Barcelona: Tusquets Editores, 1995, p. 174-175).

²⁸⁴ RAWLS, John. **Justiça e Democracia**, p. 254-255.

²⁸⁵ "Assim, de modo mais ou menos semelhante, um nome designa uma coisa, e é dado um nome a uma coisa. – Ser-nos-á freqüentemente útil se dissermos quando filosofamos: denominar algo é semelhante a colocar uma etiqueta numa coisa" (WITTGENSTEIN, Ludwig. **Investigações Filosóficas**. Tradução: José Carlos Bruni. São Paulo: Nova Cultural, 1999, p. 32).

²⁸⁶ SHANNON, Claude E. A Mathematical Theory of Communication. **The Bell System Technical Journal**, v. 27, n. 3, p. 379-423, Jul. 1948. DOI: 10.1002/j.1538-7305.1948.tb01338.x.

²⁸⁷ "This is a theory so general that one does not need to say what kinds of symbols are being considered—whether written letters or words, or musical notes, or spoken words, or symphonic music, or pictures. The theory is deep enough so that the relationships it reveals indiscriminately apply to all these and to other forms of communication" (SHANNON, Claude E; WEAVER, Warren. Op. cit., p. 25).

²⁸⁸ RAZ, Joseph. **O Conceito de Sistema Jurídico**, p. 101.

²⁸⁹ "Sendo assim, aconselhamos a falar de *normas gerais* quando estivermos diante de normas que se destinam a uma classe de pessoas; e de *normas abstratas* quando estivermos diante de normas que regulam uma ação-tipo (ou uma classe de ações. Às normas gerais se contrapõem aquelas que regulam um único indivíduo, e sugerimos que sejam chamadas de *normas individuais*; às normas abstratas se contrapõem aquelas que regulam uma ação singular, e sugerimos que sejam chamadas de *normas concretas*" (BOBBIO, Norberto. **Teoria Geral do Direito**. Tradução: Denise Agostinetti. 2ª ed. São Paulo: Martins Fontes, 2008, p. 162).

tipo geral de conduta que se aplica genericamente às pessoas²⁹⁰ e será individualizada em cada caso.²⁹¹

A desorganização é o caminho natural para o qual a informação (inclusive o Direito) caminha²⁹², e são justamente a aleatoriedade e a desorganização que marcam a flecha do tempo. Nesse sentido, a desorganização do mundo jurídico deve ser levada em consideração para a explicação dos fenômenos do Direito, para que haja percepção de que o processo entrópico levará à desorganização.

Nesse caminho, deve-se ter em vista o conceito de informação trazido por Claude Shannon, que leva em consideração que a informação se baseia no número de alternativas para que algo ocorra²⁹³, o que inclui a entropia como uma manifestação da informação como ela realmente é. Dessa maneira, o processo argumentativo deve

²⁹⁰ “Mesmo numa sociedade vasta e complexa e, como é a de um Estado moderno, há ocasiões em que uma autoridade, face a face com um indivíduo qualquer, lhe ordena que faça alguma coisa. Um policial ordena, por exemplo, a um determinado motorista que pare ou a um mendigo que se retire. Mas essas situações simples não são, nem poderiam ser a forma-padrão como o direito opera, no mínimo porque nenhuma sociedade poderia dispor do número de funcionários necessário para assegurar que cada um de seus membros fosse informado oficial e individualmente de cada ato que dele se espera. Em vez disso, essas formas particularizadas de controle ou são excepcionais, ou constituem acompanhamentos ou reforços subsidiários de formas gerais de instruções que não designam indivíduos específicos nem se dirigem a estes, e não especificam um ato em particular a ser praticado. Por isso, até a forma-padrão de uma lei penal escrita (que, entre todas as variedades do direito, é a mais semelhante a uma ordem apoiada por ameaças) é geral em dois sentidos: indica um tipo geral de conduta e se aplica a uma classe geral de pessoas, das quais se espera que estejam conscientes de que se aplica a elas e que a obedeçam”. (...)

“Portanto, o controle jurídico consiste principalmente, embora não exclusivamente, no controle exercido por meio de instruções que são *gerais* nesse duplo sentido”. (...) “Num Estado moderno compreende-se normalmente que, na ausência de indicações especiais que ampliem ou restrinjam a classe, suas leis gerais se estendem a todas as pessoas no interior de seus limites territoriais” (HART, Herbert A. L. Op. cit., p. 26-27).

²⁹¹ RAZ, Joseph. **O Conceito de Sistema Jurídico**, p. 103-104.

²⁹² “A agitação térmica é como um contínuo embaralhar de cartas de um baralho: as cartas estão em ordem, ao embaralhar se desorganizam. Assim, o calor passa do quente para o frio e não o contrário: por embaralhamento, pela desorganização natural de tudo. Ludwig Boltzmann entendeu isso. A diferença entre passado e futuro não está nas leis elementares do movimento, nem na gramática profunda da natureza. É a desorganização natural que conduz o mundo a situações cada vez menos peculiares, menos especiais” (ROVELLI, Carlo. **A Ordem do Tempo**, p. 21).

²⁹³ “A noção científica de informação foi esclarecida por Claude Shannon, matemático e engenheiro americano, em 1948, e é algo muito simples: a informação é uma medida do número de alternativas possíveis para alguma coisa. Por exemplo, se lanço um dado, este pode cair em seis faces. Se vejo que caiu em determinada face, tenho uma quantidade de informações $N = 6$, porque eram seis as possibilidades alternativas. Se não sei o dia de seu aniversário, há 365 possibilidades diferentes. Se você me diz o dia de seu aniversário, tenho uma informação $N = 365$. E assim por diante” (ROVELLI, Carlo. **A Realidade Não é o Que Parece**: a estrutura elementar das coisas. Tradução: Silvana Cobucci Leite. Rio de Janeiro: Objetiva, 2014, p. 187-188. E-book Kindle).

levar em consideração a probabilidade da transmissão de informação jurídica²⁹⁴, percebendo que, na transmissão, o ruído interpretativo é parte da entropia.²⁹⁵

Desse modo, a pretensão de correção²⁹⁶ deve também levar em conta a entropia da aplicação do direito ao considerar que há informação perdida no processo interpretativo.²⁹⁷ Portanto, o ruído interpretativo é esperado no processo de argumentação, uma vez que é um tipo de processo comunicativo. Destaque-se que, neste procedimento argumentativo, a **pretensão de correção** ganha contornos normativos e morais, uma vez que ela mesma é um elemento definidor do Direito:

“O direito é um sistema normativo que (1) formula uma pretensão de correção, (2) consiste na totalidade das normas que integram uma constituição socialmente eficaz em termos globais e que não são extremamente injustas, bem como na totalidade das normas estabelecidas em conformidade com essa constituição e que apresentam um mínimo de eficácia social ou de possibilidade de eficácia e não são extremamente injustas, e (3) ao qual pertencem os princípios e outros argumentos normativos, nos quais se apoia e/ou deve se apoiar o procedimento de aplicação do direito para satisfazer a pretensão à correção”.²⁹⁸

Dessa forma, a entropia atua num segundo nível para a pretensão de correção, pois a entropia jurídica pode ser tratada e mitigada antes e após a manifestação humana, com o auxílio de sistemas de Inteligência Artificial. Em suma, há a

²⁹⁴ “This is the probability, if the channel is in state α and symbol i is transmitted, that symbol j will be received and the channel left in state β . Thus α and β range over the possible states, i over the possible transmitted signals and j over the possible received signals. In the case where successive symbols are independently perturbed by the noise there is only one state, and the channel is described by the set of transition probabilities $p_i(j)$, the probability of transmitted symbol i being received as j ” (SHANNON, Claude E.; WEAVER, Warren. Op. cit., p. 66).

²⁹⁵ “If a noisy channel is fed by a source there are two statistical processes at work: the source and the noise. Thus there are a number of entropies that can be calculated. First there is the entropy $H(x)$ of the source or of the input to the channel (these will be equal if the transmitter is non-singular). The entropy of the output of the channel, i.e., the received signal, will be denoted by $H(y)$. In the noiseless case $H(y) = H(x)$. The joint entropy of input and output will be $H(x,y)$. Finally there are two conditional entropies $H_x(y)$ and $H_y(x)$, the entropy of the output when the input is known and conversely. Among these quantities we have the relations $H(x,y) = H(x) + H_x(y) = H(y) + H_y(x)$. All of these entropies can be measured on a per-second or a per-symbol basis” (Id.)

²⁹⁶ ALEXY, Robert. **Conceito e Validade do Direito**, p. 43.

²⁹⁷ “Evidently the proper correction to apply to the amount of information transmitted is the amount of this information which is missing in the received signal, or alternatively the uncertainty when we have received a signal of what was actually sent. From our previous discussion of entropy as a measure of uncertainty it seems reasonable to use the conditional entropy of the message, knowing the received signal, as a measure of this missing information. This is indeed the proper definition, as we shall see later. Following this idea the rate of actual transmission, R , would be obtained by subtracting from the rate of production (i.e., the entropy of the source) the average rate of conditional entropy. $R = H(x) - H_y(x)$ The conditional entropy $H_y(x)$ will, for convenience, be called the equivocation. It measures the average ambiguity of the received signal” (SHANNON, Claude E.; WEAVER, Warren. Op. cit., p. 67).

²⁹⁸ ALEXY, Robert. **Conceito e Validade do Direito**, p. 151.

possibilidade de mútua colaboração na interação humano-máquina, num duplo vetor de correção voltado ao tratamento da entropia jurídica.

Nesse sentido, um papel dos sistemas de Inteligência Artificial no processo interpretativo é o de servir como um meio comunicacional de proposições normativas²⁹⁹, previamente programadas ou inferidas a partir dos dados, de maneira a enriquecer o processo argumentativo humano.

Nessa linha, para se construir a informação jurídica, não é necessária a visão do todo (o que não é exigido pela correção procedimental relativa³⁰⁰), mas apenas das partes necessárias à construção da mensagem que se está a comunicar com o apoio do sistema de IA voltado ao auxílio à decisão judicial, podendo utilizar o modelo de Claude Shannon como uma inspiração. Assim, conforme proposto por Shannon, veja-se o diagrama esquemático³⁰¹ para um sistema de comunicação:

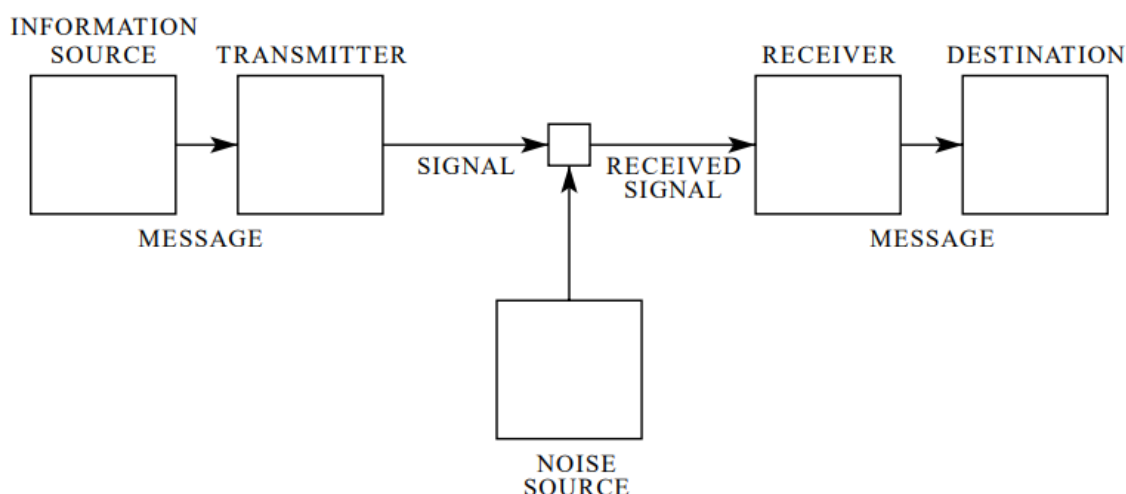


Fig. 1. – Diagrama Esquemático do Sistema Geral de Comunicação de Claude Shannon.

Dessa maneira, a pretensão de correção busca evitar a errônea aplicação do Direito e, de maneira, proativa e consciente, evitar os erros interpretativos mediante

²⁹⁹ BULYGIN, Eugenio. **Lógica Deóntica, Normas y Propositiones Normativas**. Madri: Marcial Pons, 2018, p. 60-62.

³⁰⁰ “O verdadeiro problema consiste na relativização do conceito de correção. Ele é relativizado em quatro sentidos: com vista a (1) regras do discurso, (2) medida de seu cumprimento, (3) participantes e (4) datas.

Aqui deve ser deixado de lado o primeiro problema, a relatividade com vista a *regras do discurso*. Ele não concerne aos problemas da utilidade e do status, mas ao problema da fundamentação, que não deve ser discutido aqui. Pelo menos, porém, fica claro que e como também esse problema está em uma relação para com o problema do discurso” (ALEXY, Robert. **Direito, Razão, Discurso**, p. 98).

³⁰¹ SHANNON, Claude E. *A Mathematical Theory of Communication*.

processo claro e objetivo, o que pode ser realizado através do que Claude Shannon chama de “canal de correção”.³⁰² Sobre o tema, pode ser observado o seguinte diagrama esquemático do sistema de correção proposto por Shannon³⁰³:

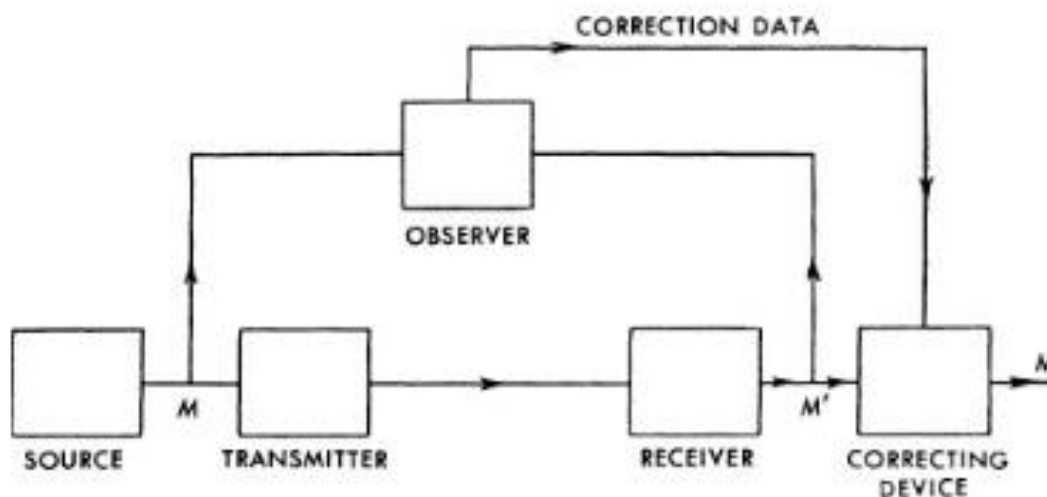


Fig. 2. – Diagrama Esquemático do Sistema de Correção de Claude Shannon.

Assim, para Claude Shannon a adequada utilização do mecanismo de correção pode suprir a grande maioria dos erros de informação, conforme o autor expôs em seu Teorema 10.³⁰⁴ Inspirando-se no pensamento de Shannon, justamente nesse aspecto, a utilização de sistemas de Inteligência Artificial para o auxílio à decisão judicial tem a possibilidade de atuar como um meio de diminuição de erros³⁰⁵ de aplicação (que ocorrem por força entrópica), atuando como um elemento de correção.

Nessa linha, passa a existir um sistema com duplo canal de correção, na medida em que a Inteligência Artificial pode atuar adequando as proposições

³⁰² “The following theorem gives a direct intuitive interpretation of the equivocation and also serves to justify it as the unique appropriate measure. We consider a communication system and an observer (or auxiliary device) who can see both what is sent and what is recovered (with errors due to noise). This observer notes the errors in the recovered message and transmits data to the receiving point over a ‘correction channel’ to enable the receiver to correct the errors” (SHANNON, Claude E.; WEAVER, Warren. Op. cit., p. 68).

³⁰³ Id.

³⁰⁴ “Theorem 10: If the correction channel has a capacity equal to $H_y(x)$ it is possible to so encode the correction data as to send it over this channel and correct all but an arbitrarily small fraction ϵ of the errors. This is not possible if the channel capacity is less than $H_y(x)$ ” (Id.).

³⁰⁵ “Theorem 11: Let a discrete channel have the capacity C and a discrete source the entropy per second H . If $H \leq C$ there exists a coding system such that the output of the source can be transmitted over the channel with an arbitrarily small frequency of errors (or an arbitrarily small equivocation). If $H > C$ it is possible to encode the source so that the equivocation is less than $H - C + \epsilon$ where ϵ is arbitrarily small. There is no method of encoding which gives an equivocation less than $H - C$ ” (Ibid., p. 71).

interpretativas humanas, enquanto o intérprete humano atua como elemento de correção de propostas de aplicação do Direito realizadas por meio computacional. Dessa forma, pela Inteligência Artificial, de um lado, há a correção da entropia da interpretação humana, e pelo intérprete humano existe a correção da entropia da sugestão computacional de aplicação do Direito, o que se dá através de “conceitos jurídicos intermediados” (“*intermediate legal concepts*” ou *ILC*)³⁰⁶ ou de fatores (*factors*) que representam atos, fatos ou estados relacionados ao suporte fático.³⁰⁷

Logo, com o avançar do grau de interação e interdependência entre ambos os canais de correção (correção da IA pelo pensamento humano e correção da decisão humana pela IA), haverá a definição, cada vez mais evidente, de que se está frente a um modelo de juiz ciborgue.

Esse duplo canal de correção implica em uma redundância que diminui a quantidade de ruído no processo de interpretação.³⁰⁸ Assim, a interação humano-máquina ocorre na medida em que o elemento humano é intrínseco ao processo de comunicação, pois se não se pode excluir da linguagem “o elemento da intenção” sem que a sua função desmorone.³⁰⁹

Dessa maneira, a menor relevância da subjetividade na implementação técnica do processo comunicacional descrito por Shannon³¹⁰ é amalgamada com a subjetividade humana³¹¹, a qual, por sua vez, é dissecada pela objetividade computacionalmente possível³¹² de um sistema de Inteligência Artificial. Isso ocorre

³⁰⁶ “*Legal rules and concepts are promulgated for normative purposes. Teleological arguments (i.e., arguments from the purposes or values served by a rule) play an important role in drawing legal analogies. Computational models that integrate legal rules, intermediate legal concepts (ILCs) from those rules, and cases applying the rules need to take underlying values into account*” (ASHLEY, Kevin D. **Artificial Intelligence and Legal Analytics**, p. 73).

³⁰⁷ BENCH-CAPON, Trevor; SARTOR, Giovanni. A Model of Legal Reasoning with Cases Incorporating Theories and Values. **Artificial Intelligence**, v. 150, n. 1-2, p. 97-143, 2003.

³⁰⁸ “*It may seem surprising that we should define a definite capacity C for a noisy channel since we can never send certain information in such a case. It is clear, however, that by sending the information in a redundant form the probability of errors can be reduced. For example, by repeating the message many times and by a statistical study of the different received versions of the message the probability of errors could be made very small*” (SHANNON, Claude E.; WEAVER, Warren. Op. cit., p. 70).

³⁰⁹ WITTGENSTEIN, Ludwig. **Observações Filosóficas**. Tradução: Adail Sobral e Maria Stela Gonçalves. São Paulo: Edições Loyola, 2005, p. 49.

³¹⁰ “*Frequently the messages have meaning; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem*” (SHANNON, Claude E. A Mathematical Theory of Communication).

³¹¹ “*Se digo que a representação tem que ter como objeto meu mundo, então não se pode dizer ‘já que, de outra maneira, eu não poderia verificá-lo’, mas ‘já, que de outra maneira, nem sequer começaria a fazer sentido para mim’*” (WITTGENSTEIN, Ludwig. **Observações Filosóficas**, p. 56).

³¹² “*The system needs to be able to identify some argument-related information in the texts, such as passages involving arguments on the topic of interest, the roles of parties in those arguments and some*

devido ao que Shannon caracteriza como “problemas semânticos”³¹³ e como “problemas de eficácia”³¹⁴, o que acaba por exigir um sistema de decodificação através de um “receptor semântico”³¹⁵ voltado à diminuição do ruído. Nesse aspecto, a função de receptor semântico pode ser realizada pelo intérprete humano.

Portanto, a visão aqui verificada tem elementos que demonstram a existência de interpretação do Direito mediante um conceito de temporalidade aleatório, voltado à desorganização, que se manifestará em erros, vieses e ruídos interpretativos que ocorrem devido ao fenômeno da entropia.

O processo de aplicação do Direito tem, em si mesmo, ao passar pelo fenômeno da individualização³¹⁶, a elaboração de uma pintura consistente na definição das proposições normativas aplicáveis³¹⁷ e na aplicação destas proposições ao suporte fático (a tela) segundo a concepção do intérprete.

Ao final do trabalho do intérprete, seja pela entropia inerente ao ruído ou pela entropia do processo humano de pensamento, haverá uma obra única consistente na aplicação do Direito e na sua demonstração através da argumentação jurídica.³¹⁸

Para que a analogia possa ficar mais evidenciada, veja-se a seguinte pintura de Claude Monet:

relevant features of the fact situation that impact the outcome of the argument such as legal factors” (ASHLEY, Kevin D. **Artificial Intelligence and Legal Analytics**, p. 166).

³¹³ *“The semantic problems are concerned with the identity, or satisfactorily close approximation, in the interpretation of meaning by the receiver, as compared with the intended meaning of the sender”* (SHANNON, Claude E; WEAVER, Warren. Op. cit., p. 4).

³¹⁴ *“The effectiveness problems are concerned with the success with which the meaning conveyed to the receiver leads to the desired conduct on his part”* (Id.).

³¹⁵ *“One can imagine, as an addition to the diagram, another box labeled “Semantic Receiver” interposed between the engineering receiver (which changes signals to messages) and the destination. This semantic receiver subjects the message to a second decoding, the demand on this one being that it must match the statistical semantic characteristics of the message to the statistical semantic capacities of the totality of receivers, or of that subset of receivers which constitute the audience one wishes to affect”* (Ibid., p. 26).

³¹⁶ RAZ, Joseph. **O Conceito de Sistema Jurídico**, p. 103-104.

³¹⁷ ALCHOURRÓN, Carlos E.; BULYGIN, Eugenio. **Normative Systems**. New York: Springer-Verlag, 1971, p. 25.

³¹⁸ “Se se deseja os conceitos de relatividade e de objetividade das normas morais, então se pode dizer que o resultado do discurso não é nem totalmente relativo nem objetivo. É relativo na medida em que se determina por meio de características dos participantes; e objetivo na medida em que depende da realização do procedimento definido através das regras do discurso. Isso significa que a comprovação discursiva não conduz à segurança, mas sim à exclusão da simples opinião e decisão, o que justifica a utilização do conceito de correção relativa” (ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 272).



Fig. 3. – *Le Jardin de l'artiste à Giverny*, de Claude Monet.³¹⁹

O impressionismo de Claude Monet traz toda a verve do artista, que pincelou na tela a representação (descrição) gráfica de um belíssimo jardim florido. Mas a pintura é uma representação parcial da realidade, o que não a invalida como um elemento de comunicação. Assim, de maneira auxiliar na formação de elementos de comunicação, pode-se verificar o alcance de um sistema de IA baseado em redes neurais de aprendizado profundo (*deep learning*), como o *DALL·E*³²⁰:



Fig. 4. – *The Artist's Garden at Giverny*, ao estilo Monet, de *DALL·E*.³²¹

³¹⁹ MONET, Claude. **Le Jardin de l'artiste à Giverny**. 1900. Óleo sobre tela. 84,6 cm x 92,6 cm.

³²⁰ OPENAI. **DALL·E**. [2022]. Disponível em: <https://labs.openai.com>. Acesso em 12 jun. 2022.

³²¹ Tendo em vista que este autor teve acesso de desenvolvedor à versão inicial do sistema (devido à pesquisa aqui desenvolvida), a imagem foi gerada pela inteligência artificial *DALL·E* em 17/07/2022, a

Uma descrição de aplicação normativa, dentre os tantos nomes pelos quais ela é chamada (proposição normativa, enunciado normativo, enunciado dogmático, conceito jurídico intermediado, etc.), não tem sequer o potencial de captar toda a realidade (o que marca a relatividade do procedimento de correção³²²). Ora, qual é a escala do mapa ou nível de detalhe das fotos de satélite de uma cidade? O mapa ou a foto de satélite, se adequadamente interpretados, têm evidente aplicabilidade para a navegação e para se conhecer Nova York, Curitiba ou Sevilha, por exemplo. Mas a foto de satélite não contém o elemento da dinâmica social nova-iorquina, a brisa fresca de Curitiba ao final da tarde ou a alegria das ruas da cidade espanhola. Toda representação é um recorte da realidade, mais ou menos arbitrário, e, com o Direito isso não é diferente.

Toda representação da realidade é parcial. Isso é destacado quando Alan Turing, no famoso artigo *On Computable Numbers*³²³, descreve a abordagem computacional dos problemas como limitada por uma finitude dos números de registros que podem ser realizados em memória. Dessa forma, a concepção de computação *Turing-complete*³²⁴ (Turing-completude), já desconsiderando os problemas inerentes à explosão combinatória, dependeria de uma capacidade

partir da entrada de texto “*The Artist’s Garden at Giverny, Monet*”, a tradução adequada de “*Le Jardin de l’artiste à Giverny*” de Claude Monet (DALL·E. **The Artist’s Garden at Giverny, Monet**. 17 jun. 2022. Imagem PNG. 1024 pixels x 1024 pixels).

³²² “Poderia achar-se que, em vista da relatividade do cumprimento, dos participantes e da data, deveria renunciar-se ao conceito de correção. Em vez disso, assim se poderia propor, deveriam ser empregados melhor conceitos como os da sustentabilidade, da plausibilidade, da fundamentalidade e da razoabilidade. Isso teria, de fato, a vantagem que não se teria mais de trabalhar com um conceito fendido de correção: com o da correção procedimental absoluta, que tem caráter de uma ideia meramente regulativa e o da correção procedimental relativa, que abarca uma medida alta de incerteza. Trata-se aqui, como tão frequentemente, de um problema terminológico, no que se refretem questões substanciais sérias. Em todo o caso, fala em favor da terminologia, aqui privilegiada, que o uso do idioma cotidiano admite, sem mais, em vista de duas concepções normativas em sentido contrário, bem fundamentadas, dizer que ambos os argumentadores *a seu modo* têm razão (correção relativa) e, então, perguntar, mais além, quem então *realmente* tem razão (correção absoluta), em que se pode, simultaneamente, conceder que certamente ninguém conseguirá averiguar isso. O fundamento substancial para a terminologia aqui escolhida é que a revisão discursiva, sem dúvida, não leva ao âmbito da segurança, contudo, para fora do âmbito do mero achar e considerar verdade. Em vista do fato que mais em questões práticas não é possível, parece-me ser conveniente o emprego do conceito de correção relativa” (ALEXY, Robert. **Direito, Razão, Discurso**, p. 99-100).

³²³ “*I assume then that the computation is carried out on one-dimensional paper, i.e. on a tape divided into squares. I shall also suppose that the number of symbols which may be printed is finite. If we were to allow an infinity of symbols, then there would be symbols differing to an arbitrarily small extent j . The effect of this restriction of the number of symbols is not very serious*” (TURING, Alan M. *On computable numbers, with an application to the Entscheidungsproblem*).

³²⁴ POURABDOLLAH, Amir; BRAILSFORD, Tim. *On the Turing Completeness of the Semantic Web. Theory of Computing Systems*, v. 56, 2014, p. 1-18. DOI: 10.1007/s00224-014-9536-x.

infinita³²⁵ de endereçamento na memória computacional, o que existe apenas em teoria³²⁶, pois não há poder computacional infinito.

Portanto, a parcialidade da representação da realidade, por si só, não impede o aproveitamento dessa representação parcial através de sistemas de Inteligência Artificial, na elaboração de uma solução cooperativa humano-máquina para a interpretação do Direito.

Isso ocorre porque o reconhecimento do dado jurídico, abaixo do nível do argumento jurídico, como o elemento basal da constituição do Direito escapa de duas armadilhas. A primeira é a armadilha de que Direito é aquilo que o Estado diz que o é. A segunda armadilha é que Direito seria uma racionalidade superior pressuposta. A realidade básica do Direito é sua paulatina desintegração, o resto são arranjos, simplificações e alegorias que têm a função pedagógica de tentar explicar o que é o Direito. Nesse caminho, a entropia jurídica não amarra o Direito a uma única forma de representação.

À semelhança da teoria de Alexy, o que se busca é a existência de um procedimento³²⁷ que tenha a pretensão de correção, o qual irá variar a depender do grau de organização que a sociedade der ao Direito, e aos seus mecanismos de criação e de interpretação.

Dessa maneira, o Direito é informação, informação jurídica que flui de um estado de organização para um estado de desorganização, estado esse que pode ser representado (parcialmente) através de sistemas de computacionais de Inteligência Artificial, o que sempre demandará a atuação humana no processo de aplicação legítima do Direito.

Nesse contexto, reconhecida a típica relatividade dos discursos reais³²⁸, de possibilidades e limites da interação entre a pretensão de correção do intérprete

³²⁵ “A menor mudança nas condições iniciais será amplificada. O sistema é instável. Uma consequência desta instabilidade, e da sensibilidade às condições iniciais que dela resulta, é que a trajetória se torna uma idealização. Com efeito, é impossível para nós prepararmos um sistema de tal sorte que possamos atribuir-lhe uma trajetória bem determinada, pois uma tal preparação ter uma precisão infinita” (PRIGOGINE, Ilya. Op. cit., p. 39).

³²⁶ “*Most actual digital computers have only a finite store. There is no theoretical difficulty in the idea of a computer with an unlimited store. Of course only a finite part can have been used at any one time. Likewise only a finite amount can have been constructed, but we can imagine more and more being added as required. Such computers have special theoretical interest and will be called infinitive capacity computers*” (TURING, Alan M. Computing Machinery and Intelligence).

³²⁷ ALEXY, Robert. **Direito, Razão, Discurso**, p. 95.

³²⁸ “Algumas regras do discurso podem somente ou ser cumpridas ou não ser cumpridas, em outras, pelo contrário, por causa do seu caráter ideal, somente um cumprimento aproximado é possível. O cumprimento é, nelas, um assunto de grau. Por esse meio, nasce o problema da relatividade com vista

humano³²⁹ e a correção proposta através de IA, o sistema de correção jurídico pode passar a ser integrado pelo duplo canal de correção, o qual, ao se colocar na posição de “observador” no sistema de correção (vide a Fig. 2), atua por meio de ciclos internos de correção. A respeito do duplo canal de correção aqui exposto, veja-se o seguinte diagrama esquemático:

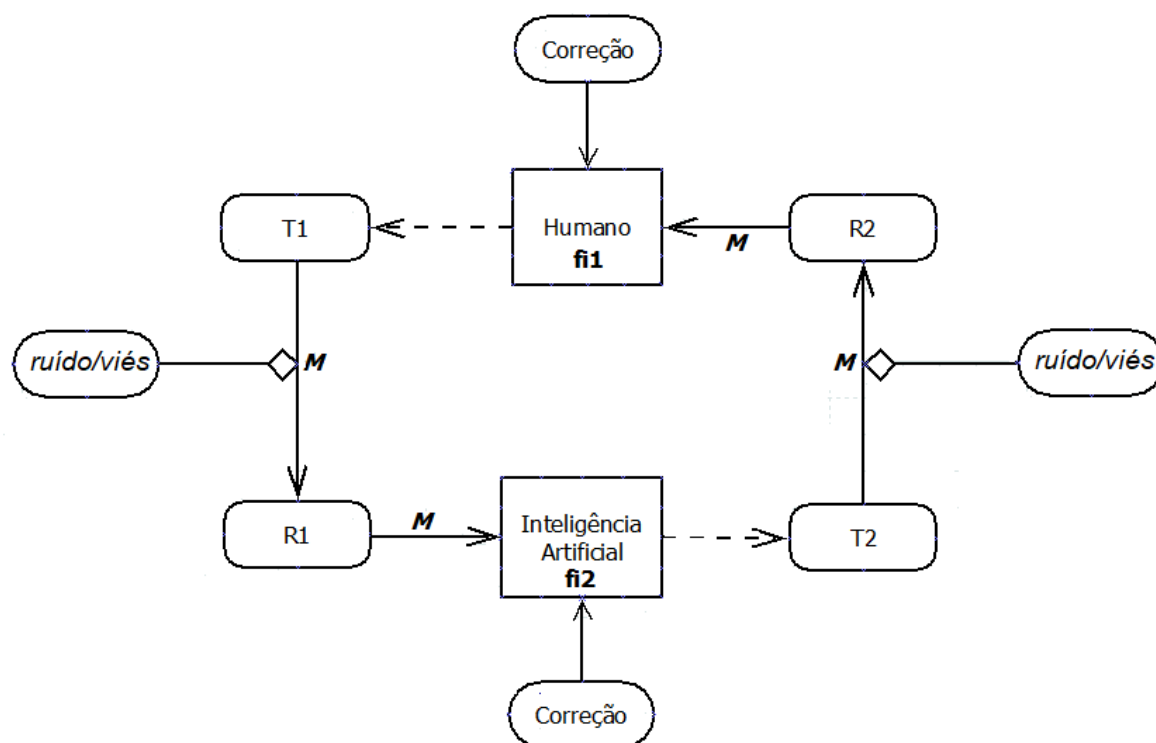


Fig. 5 – Diagrama Esquemático do Duplo Canal de Correção.

O duplo canal de correção é composto pelo elemento humano, que é a fonte de informação 1 (**fi1**), e pelo sistema de Inteligência Artificial, que consiste na fonte de informação 2 (ou **fi2**). A transmissão da mensagem (*M*) está representada em T1 e T2. A recepção da mensagem é feita por R1 e R2. Neste modelo, a entropia jurídica

à medida do cumprimento. O essencial, para isso, já dito na discussão do discurso ideal no quadro da discussão do problema do critério. A possibilidade de cumprir somente aproximativa leva, necessariamente, à incerteza do discurso como critério. Incerteza, porém, não implica imprestabilidade” (Ibid., p. 98-99).

³²⁹ “Os participantes de um discurso prático devem, independente disto, se existe uma resposta unicamente correta, promover a pretensão que sua resposta é a unicamente correta, se suas afirmações e fundamentações devem ser convenientes. Isso pressupõe somente que é possível que existam questões práticas, nas quais, no discurso, uma resposta como unicamente correta pode ser distinguida e que não é seguro quais questões são essas, de modo que vale a pena, em cada questão, tentar o descobrir de uma resposta unicamebte correta. Na base disso está uma concepção de correção *procedimental absoluta*. Ela satisfaz, sem mais, ao significado de uso corrente da expressão ‘correto’” (Ibid., p. 97).

é tratada através de ciclos no duplo canal de correção, a fim de que haja um processo comunicacional entre as proposições jurídicas humanas e as proposições normativas da IA. Logo, o ruído e os vieses, quando advindos da mensagem (*M*) que dá entrada (*input*) através do duplo canal de correção, passam por sistema híbrido argumentativo-computacional de tratamento. Da mesma forma, os ruídos e vieses, quando surgidos internamente no âmbito da proposição advinda da fonte de informação (**fi1** ou **fi2**), recebem tratamento, argumentativo (**fi1**) ou computacional (**fi2**), voltado à correção.

Dessa maneira, antes de se adentrar nos reflexos do duplo canal de correção para o delineamento do modelo de juiz ciborgue (o que será demonstrado nos próximos capítulos), algumas premissas básicas devem ser demarcadas desde já, uma vez que isso possibilitará a melhor compreensão dos demais capítulos desta obra. Assim, a descrição básica do funcionamento do processo argumentativo-comunicacional interno que ocorre no duplo canal de correção passa pelos seguintes elementos:

1. No funcionamento interno do duplo canal de correção, o destinatário da mensagem é também corretor (tem função de correção);
2. Humano e IA (Inteligência Artificial) podem buscar qualquer argumento juridicamente admissível para estímulo dos ciclos do duplo canal de correção;
3. Ao menos uma interação humano-máquina deve ser realizada através do duplo canal de correção;
4. Os ciclos de interação humano-máquina devem ser realizados até que haja uma justificativa humana que supere argumentativamente as soluções propostas pela IA, o que inclui a explicabilidade da decisão e dos motivos pelos quais as propostas da Inteligência Artificial foram aceitas ou rejeitadas;
5. Os ciclos de correção podem ser realizados até que o intérprete humano entenda pela juridicidade da proposição obtida;
6. As sugestões de aplicação do Direito em cada ciclo, humanas e computacionais, devem ser registradas e explicáveis;
7. No limite, o humano decide quando se encerram os ciclos de correção, desde que estritamente respeitadas as demais regras de funcionamento do duplo canal de correção.

Esses elementos possuem reflexos sobre a Teoria do Direito, especialmente na argumentação jurídica, os quais serão melhor definidos ao longo dos próximos capítulos.

2 PRINCIPAIS ASPECTOS DA INTELIGÊNCIA ARTIFICIAL APLICADA AO DIREITO

2.1 Inteligência Artificial Judicial: A Construção da Ideia de Explicabilidade

A inclusão do debate sobre o direito à explicabilidade (ou direito à explicação) é necessária para contextualizar a tese em um quadro mais amplo da atual fase de estudos da Inteligência Artificial, pois, para a grande maioria dos leitores da área jurídica, o tema possui questões técnicas que precisam ser enquadradas em termos que lhes sejam compreensíveis a partir de seu ponto de vista. Nesse rumo, a compreensão da tese, consistente no duplo canal de correção nascido do paradigma de juiz ciborgue, passa pela análise da explicabilidade de sistemas de Inteligência Artificial, pois o direito à explicação é parte integrante da tese. O juiz ciborgue sem dever de explicabilidade poderia se transformar em um tirano, o que exige que o tema seja abordado nesta obra.

O conhecimento não é feito só com o “velho”, pois é fruto de uma obra de desconstrução para uma reconstrução. O pensamento de Jürgen Habermas³³⁰, por exemplo, mostra isso. A ciência vive de uma arquitetura construída “sobre os ombros de gigantes”³³¹, mas, às vezes, de tempos em tempos, passa por grandes transformações. Para ilustrar isso, algumas provocações são necessárias.

Por exemplo, no filme *Matrix Reloaded*, uma obra de ficção, após Neo perceber que fundamentalmente “o problema é a escolha”, o Arquiteto explica suas tentativas de simular um mundo (e uma sociedade) aceitável para a humanidade.³³² No diálogo

³³⁰ HABERMAS, Jürgen. **Teoria do Agir Comunicativo**, 1, p. 205-206.

³³¹ “If I have seen further it is by standing on the shoulders of Giants” (NEWTON, Isaac. **Letter from Sir Isaac Newton to Robert Hooke**, 1675. Disponível em: <https://digitallibrary.hsp.org/index.php/Detail/objects/9792>. Acesso em: 06 abr. 2021).

³³² “Arquiteto: ‘— A primeira Matrix que eu projetei era evidentemente perfeita, uma obra de arte, impecável, sublime. Um triunfo equiparado apenas ao seu fracasso monumental. A inevitabilidade de sua ruína é tão evidente para mim agora quanto é uma consequência da imperfeição inerente a todo ser humano. Assim sendo, eu a redesenhei com base na história de vocês para refletir com maior precisão as variações grotescas de sua natureza. Todavia, mais uma vez, eu fui frustrado pelo fracasso. Desde então, compreendi que a resposta me escapava, porque ela necessitava de uma mente inferior, ou talvez uma mente menos afeita aos parâmetros da perfeição. Portanto, a resposta foi encontrada, por acaso, por outrem, um programa intuitivo, inicialmente criado para investigar certos aspectos da psique humana. Se eu sou o pai da Matrix, ela sem dúvida seria a mãe’.

Neo: — ‘A Oráculo!’

Arquiteto: — ‘Oh, por favor... Como eu dizia, ela se deparou por acaso com uma solução por meio da qual **quase 99,9% de todas as cobaias aceitavam o programa, contanto que lhes fosse dada uma escolha, mesmo que só estivessem cientes dela em um nível quase inconsciente.** Embora esta

longo entre *Neo* e o *Arquiteto* a tensão central se pautou no que era a justiça das máquinas para o que restava da (e de) humanidade.³³³

Para o personagem *Arquiteto* (uma IA forte³³⁴ com muitíssimos recursos computacionais para prover seu aprendizado de máquina), algumas perguntas foram sendo respondidas através da tentativa e erro. Primeiramente, foi criada para a humanidade uma simulação “perfeita” e “sublime”, a qual teve como destino o “fracasso monumental”³³⁵ decorrente de sua inaceitabilidade pelos humanos, devido ao fato de que a visão do que é um paraíso não é idêntica para seres humanos imperfeitos e diferentes entre si. A justiça maquinal também criou a simulação de um mundo infernal para “refletir com maior precisão as variações grotescas” da natureza humana³³⁶, versão pautada no sofrimento que também foi rejeitada pela humanidade. Por fim, a justiça que funcionou, segundo o *Arquiteto*, foi um subcontrato social por assentimento subliminar de humanos dominados por um poder de fato da IA, já que “quase 99,9% de todas as cobaias” aceitam a programação, desde que seja oportunizada uma escolha.³³⁷ Contudo, o simples assentimento subliminar sem explicabilidade consciente é um tipo de controle dos destinos individuais que não se amolda ao direito dos cidadãos de conhecer as causas e os motivos das decisões públicas, fazendo com que seja apenas um simulacro em forma de holograma³³⁸ jurídico-político.

A referência às obras de ficção é uma fagulha para acender o debate, um estímulo ao pensamento acerca de quais seriam os pressupostos jurídico-filosóficos para a constituição de uma cooperação humano-máquina, na busca de explicar a simbiose entre o pensamento humano e suas típicas limitações (do pensamento rápido através de heurísticas e do pensamento reflexivo)³³⁹ em interação com um aparato de soluções decisórias trazidas pelas técnicas de Inteligência Artificial.³⁴⁰

resposta funcionasse, ela era óbvia e fundamentalmente defeituosa, criando, assim, a anomalia sistêmica contraditória, a qual, sem vigilância, poderia ameaçar o próprio sistema. Por conseguinte, aqueles que recusavam o programa, ainda que uma minoria, se não vigiados, constituiriam uma probabilidade crescente de catástrofe” (THE MATRIX Reloaded. Direção: de Lana Wachowski e Lilly Wachowski. Produção: Warner Bros. EUA: 2003).

³³³ Id.

³³⁴ BOSTROM, Nick. Op. cit., posição 880.

³³⁵ THE MATRIX Reloaded. Op. cit.

³³⁶ Id.

³³⁷ Id.

³³⁸ BAUDRILLARD, Jean. **Simulacros e Simulação**. Lisboa: Relógio d'Água, 1991, p. 133-135.

³³⁹ KAHNEMAN, Daniel. Op. cit., posição 613.

³⁴⁰ TAULLI, Tom. **Introdução à Inteligência Artificial**: Uma abordagem não técnica. São Paulo: Novatec, 2020, p. 187-190.

Assim, as preocupações com aspectos da justiça numa interação humano-máquina, ainda que tenham se iniciado em obras de ficção científica, permeiam cada vez mais a realidade do mundo atual. Um exemplo disso é a predição da prática de crimes, que era reservada à ficção científica³⁴¹, e tornou-se parcial realidade com o sistema COMPAS³⁴², cujas decisões se mostraram com viés racial.³⁴³

Conforme anteriormente referido, as três principais posições perante o problema têm sido o alarmismo apocalíptico (exemplificado no pensamento de Yuval Harari³⁴⁴), o desalento com o novo (que tem tomado a psique de grande parte dos juristas³⁴⁵) e o arrebatamento tecnológico (de parte da comunidade científica³⁴⁶). Todavia, é necessário um pouco de realidade como temperamento às posições referidas.

O desvelar do novo insiste em ficar no futuro³⁴⁷, longe dos olhos curiosos do presente. Todavia, já existe muito presente para ser trabalhado. Nessa linha, um estudo realizado pela Fundação Getúlio Vargas, através de seu Centro de Inovação, Administração e Pesquisa do Judiciário, indicou que existem sessenta e quatro

³⁴¹ Por exemplo: DICK, Philip K. **The Minority Report**: And Other Classic Stories. New York: Citadel Press, 2002.

³⁴² ANGWIN, Julia et al. **Machine Bias**. Disponível em: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. Acesso em: 28 jun. 2022.

³⁴³ CHOULDECHOVA, Alexandra. Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments. **Big Data**, v. 5, n. 2, p. 153-1163, 2017. DOI: 10.1089/big.2016.0047.

³⁴⁴ HARARI, Yuval Noah. **Homo Deus**, p. 313.

³⁴⁵ *"A second common bias is 'irrational rejectionism', mentioned in the Introduction. I define this as the dogmatic dismissal of a system with which the critic has no direct personal experience. With no sense of embarrassment, lawyers frequently reject the idea of online courts having neither seen them in action nor taken the time to learn what is actually being proposed. Without reference to the problems that online courts might actually solve, appeals are immediately made to access to justice, problems and objections flow freely while operational systems are summarily dismissed. Without so much a glance at a screen. Sometimes this bias is rooted in a fear of the unknown or a sense that judges and lawyers know best when it comes to reform of the courts. Often it is based on an honest instinct that what is envisaged is neither possible nor desirable. But, there can be little doubt that an unwillingness to keep an open mind to new technologies is a serious obstacle to progress"* (SUSSKIND, Richard. **Online Courts and the Future of Justice**. Oxford: Oxford University Press, 2019, p. 44. E-book Kindle).

³⁴⁶ *"Technology companies will become the arbiters of human rights, because it will be through their devices that we will learn of abuse, and through their devices that abuse is carried out. National security agencies around the world will be engaged in conversations about cyberdeterrence that continue to use the language of nuclear deterrence developed for an earlier era. This may not seem productive, but having political leaders fear the consequences of launching a major cyberassault is likely to result in a stable balance of power maintained and monitored through the internet of things. Instead of being caught up in an arms race, political actors will race to develop better bots. With device networks expanding rapidly, botnets will have many more nodes and passages in which to grow, and many more sensors that can be activated to surveil and manipulate public opinion"* (HOWARD, Philip N. Op. cit., p. 233-234).

³⁴⁷ *"A maioria das previsões de longo alcance do que é tecnicamente factível em tempos futuros subestima dramaticamente a potência dos desenvolvimentos futuros, porque se baseiam no que eu chamo de visão 'linear intuitiva' da história mais do que na visão 'exponencial histórica'"* (KURZWEIL, Ray. **A Singularidade Está Próxima**, p. 27).

projetos³⁴⁸ de Inteligência Artificial judicial em quarenta e sete Tribunais do Brasil³⁴⁹, muitos desses projetos já estão em fase de implementação. Vários dos projetos de IA judicial, como, por exemplo, os sistemas *Victor* (STF), *Sócrates* (STJ), *Sinapse* (TJRO), *Sigma* (TRF3), *Sugestão de Modelos de Minutas* (TRF4) e *JUDI* (TJSP) impactam diretamente no processo de decisão dos julgadores, o que denota a relevância do debate acerca do presente tema.

Levando em conta essa realidade, o atual sentimento de estranhamento com a tecnologia se constrói a partir da crise com aquilo que ainda não existe (mas poderá existir), mesclado com o espanto com aquilo que já existe. Essa é a fotografia do presente momento da relação entre decisão judicial e IA. Todavia, olhares criteriosos são necessários, pois a decisão judicial no entrelaçamento juiz-IA é um fenômeno que traz riscos e potencialidades antes não experimentados. No plano dos riscos emergentes, destacam-se os vieses algorítmicos³⁵⁰, que têm atraído bastante atenção da comunidade jurídica (embora os vieses já habitassem a mente dos julgadores muito antes dos computadores existirem³⁵¹), e a maior opacidade das decisões auxiliadas por IA, sendo esse último o ponto de tensão no qual o presente tópico se foca.

Todavia, para ter condições de escolher, o ser humano deve efetivamente conhecer as decisões que são tomadas na sociedade democrática³⁵² da qual ele faz

³⁴⁸ Os principais são: *Victor* no STF; *Athos*, *Sócrates*, *E-Juris* e *TUA* no STJ; *BEM-TE-VI* no TST; Plataforma *Sinapses* no CNJ; *Robô Secor*, *Banco de Sentenças*, *SIB*, *ALEI* e *Projeto Execução Célere* no TRF1; *Atendente Virtual* no TRF2; *SINARA*, *SIGMA* e *Prevenção* no TRF3; *Classificação de Temas na Vice-Presidência* e *Turmas Recursais*, *Análise de Assunto dos Processos*, *Triagem Automática de Processos a Partir da Petição Inicial* e *Sugestão de Modelos de Minutas* no TRF4, *JULIA* no TRF5; *LEIA* em seis Tribunais de Justiça diversos (TJAC, TJAL, TJAM, TJCE, TJMS e TJSP); *Hércules* no TJAL; *Queixa Cidadã* no TJBA; *Hórus* e *Ámon* no TJDF; *IA332* no TJGO; *ELIS* no TJPE; *Sinapse* no TJRO; *SCRIBA* e *Mandamus* no TJRR; *JUDI* no TJSP; *MINERJUS* no TJTO; *Inteligência Artificial e Eficiência do Judiciário* no TRT1; *Clusterização de Processos* no TRT4; *GEMINI* em três Tribunais Regionais do Trabalho diversos (TRT5, TRT15 e TRT17); *B.I. TRT 11* no TRT11; e *CONCILIA JT* no TRT12.

³⁴⁹ SALOMÃO, Luis Felipe (coord.) et al. Op. cit.

³⁵⁰ YAPO, Adrienne; WEISS, Joseph. Op. cit.

³⁵¹ SUNSTEIN, Cass R. Algorithms, Correcting Biases.

³⁵² "Una constitución puede entenderse como un proyecto histórico que los ciudadanos prosiguen de nuevo en cada generación. En el estado democrático de derecho, el ejercicio del poder político está codificado por partida doble: el tratamiento institucionalizado de los problemas planteados y la mediación regulada procedimentalmente de los correspondientes intereses deben poder entenderse al mismo tiempo como la realización de un sistema de derechos. En la escena política, sin embargo, se encuentran enfrentados los actores políticos que discuten sobre fines colectivos y sobre la distribución de los bienes colectivos. Sólo ante los tribunales y en el discurso jurídico se ponen en juego los derechos individuales reclamables judicialmente. También el derecho vigente debe interpretarse de una manera nueva en los contextos cambiantes a la vista de las nuevas necesidades y constelaciones de intereses" (HABERMAS, Jürgen. *La Lucha por el Reconocimiento en el Estado Democrático de Derecho*. In: TAYLOR, Charles; et al. **El Multiculturalismo y "La Política del Reconocimiento"**.

parte³⁵³, o que inclui as decisões auxiliadas por sistemas de IA judicial. Dessa forma, o desenvolvimento de Inteligência Artificial explicável (xAI) é necessário para que os padrões democráticos pautados na transparência das decisões públicas sejam respeitados.

Nessa linha, a Inteligência Artificial explicável (xAI) envolve esforços para explicar como um modelo de IA chegou ao resultado (*output*) verificado³⁵⁴ ou ainda para auxiliar os humanos na interpretação desse.³⁵⁵

Em relação às decisões judiciais auxiliadas por Inteligência Artificial, de maneira inerente à arena deliberativa, a pretensão de validade nas decisões de repercussão na sociedade foi trazida por Jürgen Habermas, pretensão esta que exige a exteriorização dos meios de decisão e dos motivos pelos quais a decisão é tomada, haja vista o necessário controle procedimental da verdade proposicional, da correção normativa e dos aspectos subjetivos das decisões públicas.³⁵⁶ Essa abordagem é

Tradução: Mónica Utrilla de Neira, Liliana Andrade Llanas e Gerard Vilar Roca. Ciudad de México: Fondo de Cultura Económica, 2009, p. 156. E-book Kindle).

³⁵³ “Dado que nos colocamos na tradição do pensamento democrático, devemos também considerar os cidadãos como pessoas livres e iguais. A idéia intuitiva básica, aqui, é a de que, em virtude das suas capacidades morais, assim como das da razão, do pensamento e do julgamento que a ela estão ligadas, dizemos que essas pessoas são livres.” (...) “O senso da justiça é a capacidade de compreender, aplicar e respeitar nos seus atos a concepção pública da justiça que caracteriza os termos de uma cooperação equitativa. E ser capaz de uma concepção do bem é poder formar, revisar e buscar racionalmente uma concepção de nossa vantagem ou bem. No caso da cooperação social, é preciso não tomar esse bem no sentido estreito, mas concebê-lo como tudo o que tem valor na vida humana” (RAWLS, John. **Justiça e Democracia**, p. 216).

³⁵⁴ “xAI encompasses a range of efforts to explain—or help humans interpret—how a particular machine learning model reached its conclusion. The concept of an explanation here ‘has come to refer to providing insight into the internal state of an algorithm, or to human-understandable approximations of the algorithm.’ xAI provides a variety of benefits: It can foster trust between humans and the system, identify cases in which the system appears to be biased or unfair, and bolster our own knowledge of how the world works. As discussed below, in legal settings xAI can benefit judges who wish to rely on the algorithms for decisional support, litigants who seek to persuade judges that their use of algorithms is defensible, and defendants who wish to challenge predictions about their dangerousness. xAI is not without costs, however. Most significantly, making an algorithm explainable may result in a decrease in its accuracy. xAI may also stifle innovation, force developers to reveal trade secrets, and impose high monetary costs because xAI can be expensive to build” (DEEKS, Ashley. The Judicial Demand for Explainable Artificial Intelligence. **Columbia Law Review**, v. 119, n. 7, p. 1829-1850, 2020, p. 1834).

³⁵⁵ “As a first step towards creating explanation mechanisms, there is a new line of research in interpretability, loosely defined as the science of comprehending what a model did (or might have done). Interpretable models and learning methods show great promise; examples include visual cues to find the ‘focus’ of deep neural networks in image recognition and proxy methods to simplify the output of complex systems. However, there is ample room for improvement, since identifying dominant classifiers and simplifying the problem space does not solve all possible problems associated with understanding opaque models” (GILPIN, Leilani H. et. al. Explaining Explanations: An Overview of Interpretability of Machine Learning. **2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)**, Turin, Italy, p. 80-89, 2018. DOI: 10.1109/DSAA.2018.00018).

³⁵⁶ “O conceito de entendimento remete a um comum acordo almejado pelos participantes e racionalmente motivado, que se mede segundo pretensões de validade criticáveis. A pretensão de validade (verdade proposicional, correção normativa e verdade subjetiva) caracterizam diferentes

ampliativa em relação à exigência de explicabilidade para as decisões de impacto na sociedade, embora lembre, em certos aspectos, a concepção de razão pública advinda da ideia de justiça política típica da segunda fase pensamento de John Rawls.³⁵⁷ A concepção de pretensão de validade traz consigo uma abertura a diversas possibilidades do que pode ser considerado adequado para a sociedade. Tais possibilidades são construídas no âmbito da deliberação democrática³⁵⁸, voltada a um sistema constitucional que se pauta na explicabilidade das decisões, que devem ser “razoáveis”³⁵⁹ e ter o pluralismo como característica³⁶⁰ que inspira o procedimento de deliberação.

categorias de um saber que se corporifica simbolicamente em exteriorizações. Essas exteriorizações podem ser analisadas mais de perto: por um lado, sob o aspecto da possibilidade de fundamentar exteriorizações como essas; por outro, sob o aspecto de como os atores se relacionam, por meio delas, com alguma coisa no mundo. O conceito de racionalidade comunicativa remete a diversas formas de resgate discursivo das pretensões de validade (daí Wellmer falar também de uma racionalidade ‘discursiva’) e, também, a referências de mundo aceitáveis pelas pessoas que agem comunicativamente, à medida que manifestam pretensões de validade para suas exteriorizações (e por isso a descentração da compreensão do mundo revelou-se como a dimensão mais importante da formação de uma compreensão do mundo)” (HABERMAS, Jürgen. **Teoria do Agir Comunicativo**, 1, p. 147-148).

³⁵⁷ GARGARELLA, Roberto. **As Teorias da Justiça Depois de Rawls**: Um breve manual de filosofia política. Tradução: Alonso Reis Freire. São Paulo: WMF Martins Fontes, 2008, p. 227-228.

³⁵⁸ “A justificativa da existência do Estado não reside primariamente na proteção de direitos subjetivos iguais, e sim na garantia de processo inclusivo de formação da opinião e da vontade, dentro do qual civis livres e iguais se entendem sobre quais normas e fins estão no interesse comum de todos. Com isso, exige-se do cidadão republicano muito mais do que a simples orientação pelo próprio interesse” (HABERMAS, Jürgen. **Direito e Democracia**, p. 335).

³⁵⁹ “Uma justificação, quando se trata de justiça política, se dirige aos ‘outros’, àqueles que estão em desacordo ‘conosco’, e, por esse fato, emana de um certo consenso, de hipóteses que os outros e nós mesmos admitimos como verdadeiras ou suficientemente razoáveis quando se busca alcançar um acordo realista referente aos fundamentos da justiça política. Dado o ‘fato do pluralismo’ e o fato de que uma justificação deve emanar de um consenso inquestionável, nenhuma doutrina geral e abrangente pode assumir a função básica publicamente aceitável da justiça política” (RAWLS, John. **Justiça e Democracia**, p. 254-255.).

³⁶⁰ “A existência de uma pluralidade de doutrinas abrangentes incompatíveis entre si - o fato do pluralismo - prova que a ideia de uma sociedade bem ordenada e governada pela teoria da justiça como equidade, no sentido de TJ, não é realista. Ela é incompatível com a própria realização de seus princípios nas melhores condições imagináveis. É por isso que a análise da estabilidade de uma sociedade bem-ordenada que eu propunha na terceira parte de TJ é igualmente irrealista e deve ser reformulada.

Em conclusão, o problema do liberalismo político é saber como uma sociedade democrática estável é justa, composta por cidadãos livres e iguais, mas profundamente divididos por doutrinas - religiosas, filosóficas e morais – incompatíveis entre si, pode existir de maneira durável. Dito de outra forma, como é possível que doutrinas profundamente opostas, cada uma delas querendo ser abrangente, coexistam e apoiem a concepção política representada pela democracia constitucional? Quais são a estrutura e o conteúdo de uma concepção política capaz de obter o apoio de tal consenso por justaposição?

Essa concepção política da justiça não pressupõe nenhuma doutrina abrangente particular. É por isso que ela funciona como um componente - um módulo, poder-se-ia dizer – que se pode acrescentar ou adaptar a numerosas doutrinas distintas ou que delas se pode derivar. Dessa forma, ela pode ser a base para um consenso, proveniente de uma superposição de doutrinas, em favor das instituições democráticas” (Ibid. Prefácio, p. X e XI).

É justamente ao ter por base o natural pluralismo³⁶¹ que, numa sociedade democrática, a explicabilidade das decisões tomadas por autoridades estatais ganha a importância de ser o meio apto a dar a possibilidade de publicidade, de transparência e, mais do que isso, de *accountability*³⁶² sobre referidas decisões.

Especificamente em relação ao Poder Judiciário, com a recente inclusão dos sistemas de IA judicial como novos atores no auxílio à tomada de decisão, a necessidade de explicabilidade toma corpo. A explicabilidade é trazida como um elemento da transparência das decisões tomadas por IA através de diversas iniciativas, como na Recomendação 449 da OCDE³⁶³, na *Recommendation on the Ethics of Artificial Intelligence*³⁶⁴ da UNESCO, nos *Ethical Guidelines for Trustworthy AI*³⁶⁵ da União Europeia, na previsão do direito fundamental à proteção de dados (art. 5º, LXXIX, da Constituição Federal), no art. 20 da Lei Geral de Proteção de Dados (Lei 13.709/2018) e na Resolução 332/2020 do Conselho Nacional de Justiça.³⁶⁶

O art. 19 da Resolução 332/2020 do CNJ prevê que os sistemas de Inteligência Artificial utilizados em auxílio à elaboração de decisões judiciais “observarão, como critério preponderante para definir a técnica utilizada, a explicação dos passos que conduziram ao resultado”.³⁶⁷ A transparência no uso da IA judicial, segundo o art. 8º, VI, da Resolução 332/2020 do CNJ, tem como um de seus elementos “o fornecimento de explicação satisfatória e passível de auditoria por autoridade humana quanto a qualquer proposta de decisão apresentada pelo modelo de Inteligência Artificial”.³⁶⁸ A norma do CNJ desenvolveu sua ideia de explicabilidade com inspiração aparente no **princípio da transparência judicial** previsto nos *Asilomar AI Principles*³⁶⁹, segundo o qual “qualquer envolvimento de um sistema autônomo na tomada de decisões

³⁶¹ MALISKA, Marcos Augusto. Op. cit., , p. 55.

³⁶² WAGNER, Ben. Accountability by design in technology research. **Computer Law & Security Review**, v. 37, 105398, ISSN 0267-3649, Oxford, 2020. DOI: 10.1016/j.clsr.2020.105398.

³⁶³ OCDE. **Recommendation 0449, of 21 May 2019: Recommendation of the Council on Artificial Intelligence**. Disponível em: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>. Acesso em: 04 ago. 2020.

³⁶⁴ UNESCO. Op. cit.

³⁶⁵ UNIÃO EUROPEIA. **Ethics Guidelines for Trustworthy AI**. [2019]. Disponível em: https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419. Acesso em: 25 jul. 2020.

³⁶⁶ CONSELHO NACIONAL DE JUSTIÇA. Resolução 332, de 21 de agosto de 2020.

³⁶⁷ Id.

³⁶⁸ Id.

³⁶⁹ Na *Asilomar Conference on Beneficial AI* (2017) foram definidos os vinte e três *Asilomar AI Principles*, encontro científico este que teve a participação de membros do *Future of Life Institute* (que tem dentre seus fundadores Max Tegmark, Stephen Hawking, Erik Brynjolfsson, Nick Bostrom, Elon Musk, Stuart Russell, Alan Guth e Francesca Rossi) e de outros cientistas e personalidades do mundo da tecnologia.

judiciais deve fornecer uma explicação satisfatória, passível de auditoria por uma autoridade humana competente”.³⁷⁰

As novas abordagens referidas vêm ao encontro da exigência de explicabilidade nas decisões judiciais que tradicionalmente permeia a exposição da motivação decisória, resumidamente composta pelos motivos de fato e pelos motivos de direito (normas, posicionamentos doutrinários, precedentes judiciais e sua interpretação). Isso passa pela análise das premissas utilizadas na decisão, na forma exigida no art. 489 do Código de Processo Civil e no art. 93, IX, da Constituição Federal, sendo que tais premissas advêm de uma análise mais aprofundada dos pressupostos consensuais³⁷¹ que constituem a decisão. Nessa linha, mediante processo de atribuição³⁷² pode-se inferir um direito fundamental à explicabilidade das decisões judiciais auxiliadas por Inteligência Artificial.

Contudo, a explicabilidade dos modelos de Inteligência Artificial está cercada de dificuldades técnicas que surgem desde os modelos algorítmicos utilizados na programação, como as abordagens procedimental³⁷³, de representação do conhecimento³⁷⁴ e de aprendizado de máquina (*machine learning*)³⁷⁵ que podem ser implementadas.

Muito daquilo que na atualidade se entende como simples programação (programação orientada a objetos, por exemplo) já foi tido, num passado recente, como “Inteligência Artificial”. IA não “é mágica ou ficção científica, mas ciência, engenharia e matemática”³⁷⁶. É curioso como os juristas querem definir tudo, inclusive o que é (ou não é) Inteligência Artificial, em que pese o conceito de Inteligência Artificial dependa da existência de grau de autonomia decisória inerente aos programas de computador³⁷⁷, grau de autonomia que pode ser maior ou menor.

³⁷⁰ “8) *Judicial Transparency: Any involvement by an autonomous system in judicial decision-making should provide a satisfactory explanation auditable by a competent human authority*” (FUTURE OF LIFE INSTITUTE. **Asilomar AI Principles**. [2017]. Disponível em: <https://futureoflife.org/ai-principles>. Acesso em: 05 ago. 2020).

³⁷¹ “Exteriorizações ligadas a pretensões de correção normativa e pretensões de veracidade subjetiva, similarmente ao que se dá com outros atos que se vinculam à pretensão de verdade e eficiência proposicional, cumprem o pressuposto central de racionalidade: elas podem ser fundamentadas e criticadas” (HABERMAS, Jürgen. **Teoria do Agir Comunicativo**, 1, p. 45).

³⁷² ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 74.

³⁷³ ROVER, Aires José. **Informática no Direito: Inteligência Artificial**. Curitiba: Juruá, 2001, p. 66-67.

³⁷⁴ ASHLEY, Kevin. **Artificial Intelligence and Legal Analytics**, p. 81.

³⁷⁵ LAGE, Fernanda de Carvalho. Op. cit., p. 70-71.

³⁷⁶ RUSSELL, Stuart; NORVIG, Peter. Op. cit., p. 27.

³⁷⁷ DEWEY, Daniel; RUSSEL, Stuart; TEGMARK, Max. Op. cit.

Nos sistemas de representação do conhecimento destaca-se na área de programação jurídica o desenvolvimento de sistemas especialistas³⁷⁸, bem como tem cada vez mais ganhado espaço a aplicação de sistemas baseados em aprendizado de máquina (*machine learning*).³⁷⁹

As redes neurais de aprendizado profundo, modelos que simulam uma rede extremamente conectada de neurônios em diversas camadas (*layers*), têm vantagens e desvantagens.³⁸⁰ O maior exemplo do desenvolvimento de modelos *deep learning* é a *GPT-3 (Generative Pre-training Transformer 3)*³⁸¹ da *OpenAI*³⁸², uma IA de processamento de linguagem natural (*NLP*), com pretensão de aplicabilidade geral (*task agnostic*), baseada em pesquisas em redes neurais profundas³⁸³. Os dados a respeito do desenvolvimento do *GPT-3* foram publicados pela desenvolvedora do sistema³⁸⁴, o que possibilitou a pesquisa. Em 2022, a versão *GPT-3.5* foi incorporada ao serviço *ChatGPT* para possibilitar o acesso mais intuitivo ao modelo de IA através do terminal de um *chatbot*. Posteriormente³⁸⁵, a *OpenAI* lançou novo modelo de IA, ainda mais evoluído³⁸⁶ em termos de poder computacional e de uso do Processamento de Linguagem Natural (PLN)³⁸⁷, que é denominado *GPT-4*

³⁷⁸ ROVER, Aires José. **Informática no Direito**, p. 217-220

³⁷⁹ ROVER, Aires José. *Machine Learning* no Poder Judiciário: uma biblioteca temática, ano 2020. In: HARTMANN PEIXOTO, Fabiano (Org.). **Inteligência Artificial: estudos de inteligência artificial**. Vol. 4. Curitiba: Alteridade, 2021, p. 20-21.

³⁸⁰ MIT Technology Review. **Por que a GPT-3 é o melhor e o pior da IA atualmente**. [2021]. Disponível em: <https://mittechreview.com.br/por-que-a-gpt-3-e-o-melhor-e-o-pior-da-ia-atualmente>. Acesso em: 08 abr. 2021.

³⁸¹ BROWN, Tom B. et al. **Language Models are Few-Shot Learners**. OpenAI, San Francisco, arXiv:2005.14165. Disponível em: <https://arxiv.org/pdf/2005.14165.pdf>. Acesso em: 05 abr. 2021.

³⁸² OPENAI. **OpenAI API**. [2021]. Disponível em: <https://openai.com/blog/openai-api/>. Acesso em 08 abr. 2021.

³⁸³ Desde o lançamento do *GPT-3*, em 2020, houve grande impacto na comunidade científica. O *GPT-3* é uma IA de processamento de linguagem natural (*NLP*). O *GPT-3* é uma IA *task agnostic* via rede neural profunda (*DNN*) com 175 bilhões de parâmetros. O treinamento do sistema *deep learning* usou o seguinte *dataset* em número de *tokens*: *Common Crawl* (410 bilhões – peso 60%), *WebText2* (19 bilhões), *Books1* (12 bilhões), *Books2* (55 bilhões) e *Wikipedia* (3 bilhões). A maior base de dados é o *Common Crawl*, que é um *crawler* de leitura de conteúdos existentes em toda a internet. Em relação à língua dos textos utilizados no treinamento tem a seguinte distribuição por número de palavras: a) inglês – 92,64708% – 181.014.683.608; b) francês – 1,81853% – 3.553.061.536; c) alemão – 1,46937% – 2.870.869.396; d) espanhol – 0,77289% – 1.510.070.974; e) italiano – 0,60793% – 1.187.784.217; f) português – 0,52483% – 1.025.413.869; e outros 112 idiomas em menores percentuais.

³⁸⁴ BROWN, Tom B. et al. Op. cit.

³⁸⁵ Em março de 2023.

³⁸⁶ Maiores detalhes do *GPT-4* serão explorados no subcapítulo 5.8.

³⁸⁷ O tema do Processamento de Linguagem Natural (PLN) será tratado no subcapítulo 5.8.

(*Generative Pre-training Transformer 4*) e que possui maior criatividade, possibilidade de entradas (*inputs*) visuais e melhor compreensão do contexto de utilização.³⁸⁸

A API do GPT-3.5 e do GPT-4 já tem sido aplicada em problemas jurídicos conjuntamente com outros modelos de IA, como no caso da CARA A.I. da empresa Casetext³⁸⁹, por exemplo, um sistema hibridizado (sistema da OpenAI e IA com desenvolvimento para tarefas específicas) que é utilizado para pesquisa de jurisprudência, leis e doutrina, bem como é capaz de gerar texto com o emprego de conceitos jurídicos, o que tem auxiliado os profissionais do Direito.³⁹⁰

Contudo, o desenvolvimento de sistemas baseados em aprendizado profundo (*deep learning*), embora seja de grande valia no auxílio a advogados, por exemplo, traz consigo riscos³⁹¹, como a opacidade³⁹², caso tais sistemas sejam utilizados para auxiliar na tomada de decisões judiciais, uma vez que existem grandes dificuldades técnicas em relação à possibilidade de explicabilidade das saídas (*outputs*) dessa classe de modelos de IA³⁹³, o que impacta na transparência inerente aos padrões mínimos de justiça que se busca desenvolver numa sociedade democrática.

Sendo assim, tendo por base as premissas adotadas e os desafios verificados, pode-se passar ao próximo tópico.

³⁸⁸ OPENAI. **GPT-4 can solve difficult problems with greater accuracy, thanks to its broader general knowledge and problem solving abilities.** [2023]. Disponível em: <https://openai.com/product/gpt-4>. Acesso em: 20 mar. 2023.

³⁸⁹ CASETEXT. **CARA A.I.** [2021]. Disponível em: <https://casetext.com/cara-ai>. Acesso em: 28 abr. 2021.

³⁹⁰ Não há ainda uma visão de substituição dos operadores do Direito pela IA referida, pois esta tem menor acurácia que os humanos na atribuição de significado, uma vez que a acurácia é de 76% no GPT-3 Small e de 52% no GPT-3. A equipe de pesquisa da OpenAI, mediante os dados de desempenho fornecidos, admite que o modelo completo (GPT-3) com 175 bilhões de parâmetros tem baixa acurácia na tarefa de interpretar o real significado das frases geradas (BROWN, Tom B. et al. Op. cit.).

³⁹¹ O uso potencialmente inadequado do GPT-3 foi noticiado, por exemplo, em relação ao chamado “Project December”, no qual a API do sistema de processamento de linguagem natural foi utilizada em um aplicativo que simulava diálogos do desenvolvedor com sua falecida noiva. Referido caso suscitou diversos questionamentos éticos e teve de ser descontinuado pela OpenAI. Para maiores detalhes sobre o caso: QUACH, Katyanna. **A developer built an AI chatbot using GPT-3 that helped a man speak again to his late fiancée. OpenAI shut it down.** [2021]. Disponível em: https://www.theregister.com/2021/09/08/project_december_openai_gpt_3/. Acesso em 03 jan. 2023.

³⁹² “Machine learning tools hold out the possibility of making more accurate decisions, faster, based on far larger quantities of data than humans can process and manipulate. Importantly, though, because a machine learning system learns on its own and adjusts its parameters in ways its programmers do not specifically dictate, it often remains unclear precisely how the system reaches its predictions or recommendations. This is particularly true for ‘deep learning’ systems that use ‘neural networks,’ which are intended to replicate neural processes in the human brain” (DEEKS, Ashley. Op. cit., p. 1829).

³⁹³ PASQUALE, Frank. **The Black Box Society: The Secret Algorithms That Control Money and Information.** Cambridge: Harvard University Press, 2015, p. 216-217. E-book Kindle.

2.2 Caminhos do Direito à Explicabilidade: A Objetividade Possível para a Explicabilidade de Decisões Judiciais Auxiliadas pela IA

O Juiz é uma figura demasiadamente humana, para parafrasear Nietzsche.³⁹⁴ Ele vivencia os sentimentos do mundo e isso causa influência em seus julgamentos.³⁹⁵

Quando se pensa na ideia de absoluta objetividade, destaca-se a ficção³⁹⁶ trazida na alegoria do véu da ignorância da teoria de John Rawls³⁹⁷, o que resultou numa visão de objetividade dos criadores do Estado democrático em decorrência de seu desconhecimento da realidade e de qual seria sua futura posição na sociedade. Trata-se da chamada “‘posição original’ que é uma situação puramente hipotética” [tradução nossa]³⁹⁸, como destaca o próprio John Rawls. Dessa maneira, a objetividade total exigiria o distanciamento psíquico, humano, social e antropológico dos fatos que se inter-relacionam com o mundo normativo através dos valores de cada julgador.

Todavia, ao tratar do chamado “problema da objetividade” na interpretação do Direito e sua necessária relação com a pretensão de correção, Alexy destaca a tensão entre os conceitos de correção (que depende de objetividade) e de consenso (que possui um elemento de subjetividade entre os participantes do discurso jurídico).³⁹⁹ Essa tensão é mediada pelo reconhecimento de que a prática dos julgamentos se faz através da inafastável imperfeição do “discurso real”⁴⁰⁰, no qual os participantes do

³⁹⁴ NIETZSCHE, Friedrich. **Humano, Demasiado Humano**. Tradução: Paulo César de Souza. São Paulo: Companhia das Letras, 2000, p. 4. E-book Kindle.

³⁹⁵ “As razões da justiça podem diferir, para usar uma das expressões de Smith, das razões do “amor-próprio”, e também das razões da prudência, mas as razões da justiça continuam constituindo um âmbito extremamente vasto” (SEN, Amartya. Op. cit., posição 1172).

³⁹⁶ Habermas propõe como meio de superação da armadilha conceitual da teoria de Rawls a adoção de um “procedimento aberto de uma práxis argumentativa que acate as severas pressuposições do ‘uso público da razão’ e que não descarte já de antemão o pluralismo das convicções e cosmovisões” (HABERMAS, Jürgen. **A Inclusão do Outro**: Estudos de teoria política. São Paulo: Loyola, 2002, p. 73).

³⁹⁷ RAWLS, John. **A Theory of Justice**. Cambridge: Harvard University Press, 1971, p. 136-141.

³⁹⁸ “*It is clear, then, that the original position is a purely hypothetical situation*” (Ibid., p. 120).

³⁹⁹ “Mais importante é o segundo subproblema do problema da correção, o problema da objetividade. Nele, trata-se da objeção que o enlace, feito pela teoria do discurso, dos conceitos de correção e verdade com os do discurso e do consenso inadmissível. Correção e verdade são algo objetivo. Discursos e consensos, pelo contrário, ao âmbito da conicção subjetiva e do mero achar e aceitar. A teoria do discurso não distingue suficientemente entre considerar verdade e o ser verdade. Que todos aprovam uma proposição não significa que ela é correta ou verdadeira, porque todos poderiam equivocar-se. Isso vale também para o resultado de discursos ideais” (ALEXY, Robert. **Direito, Razão, Discurso**, p. 94-95).

⁴⁰⁰ “A quarta relatividade, a relatividade com vistas a *datas*, é inevitável perante os limites de cada discurso real. Em resultados só discursivamente possíveis, portanto, dissensos, a ideia regulativa da correção exige o caráter não definitivo do resultado. Em resultados discursivamente necessários,

discurso, embora possam ser discordantes em relação aos argumentos racionais⁴⁰¹, possuem uma “capacidade de juízo” para a diferenciação de fundamentos bons em relação aos fundamentos ruins.⁴⁰² Essa é uma das razões pela qual “não é acertado imputar à teoria do discurso que ela considera o consenso como *fundamento* para a correção”⁴⁰³, pois é a realização do procedimento argumentativo que traz consigo a objetividade que leva à correção na interpretação do Direito.⁴⁰⁴ Note-se que a objetividade, inerente ao procedimento voltado à correção, pode ser aprimorada através da inclusão de sistemas de Inteligência Artificial no auxílio às decisões judiciais. Desse modo, mesmo que o objetivo seja uma correção procedimental absoluta, que é muito difícil de ser alcançada, chega-se a uma correção procedimental relativa⁴⁰⁵, que é possibilitada pela objetividade do processo discursivo.

Dessa maneira, a objetividade advinda da pretensão de correção atua por meio de “ligação racionalmente motivada”⁴⁰⁶ entre os participantes do discurso jurídico, a qual ganha alcance na compreensão situacional em um processo de decisão que, embora não seja perfeito, busca a sua correção por todos os meios e instrumentos possíveis, o que inclui sistemas de auxílio à decisão judicial baseados em Inteligência Artificial e que contribuem para a explicabilidade do que é decidido.

Noutro enfoque, veja-se que nem mesmo o criticado “véu da ignorância” atrelado à teoria da posição original de John Rawls se volta à pretensão de uma

portanto, consensos, o caráter necessariamente provisório resulta da imperfeição necessária de cada discurso real” (Ibid., p. 99).

⁴⁰¹ “A teoria do discurso pressupõe que os participantes do discurso, portanto, pessoas, assim como elas faticamente existem, em princípio, estão capacitados a distinguir fundamentos bons de ruins para declarações substanciais. Ela parte, portanto, de uma capacidade de juízo, em princípio existente, suficiente dos participantes. Isso não significa que uma capacidade de juízo suficiente é uma exigência do procedimento. A relação entre o procedimento do discurso e a capacidade de juízo suficiente de seus participantes corresponde antes à entre a constituição de um estado constitucional democrático e a capacidade de seus cidadãos para atividades políticas, econômicas e sociais. O último não é exigido por normas constitucionais, mas pressuposto pela constituição. Acresce a isso, que é pressuposto somente uma capacidade de juízo, *em princípio* existente, suficiente. É uma das finalidades do procedimento do discurso levá-la ao desenvolvimento” (Ibid., p. 95).

⁴⁰² Ibid., p. 95-96.

⁴⁰³ Ibid., p. 95.

⁴⁰⁴ “Não o consenso, mas a realização do procedimento segundo as regras do discurso é o critério de correção verdadeiro da teoria do discurso. A questão decisiva diz o que a realização de um procedimento segundo as regras do discurso, no essencial, formais tem a ver com a correção quanto ao conteúdo das declarações normativas substanciais” (Id.).

⁴⁰⁵ Ibid., p. 99-100.

⁴⁰⁶ “Com a força ilocucionária de uma externalização, o falante pode motivar o ouvinte a aceitar sua oferta de um ato de fala e, com isso, motivá-lo a *estabelecer uma ligação racionalmente motivada*. Tal concepção prevê que sujeitos aptos a falar e agir possam fazer referência a mais que um único mundo; e que, ao se endenderem uns com os outros sobre alguma coisa em um mundum único, embasem sua comunicação sobre um sistema de mundos que suponham de maneira compartilhada” (HABERMAS, Jürgen. **Teoria do Agir Comunicativo**, 1, p. 484).

objetividade absoluta, pois a primeira fase do pensamento de Rawls busca conciliar a proteção às liberdades individuais ao caráter institucional da justiça⁴⁰⁷, já que a sociedade nascida do consenso do “véu da ignorância” estaria moldada institucionalmente por meio de valores definidos, restringindo as possibilidades de escolha das pessoas que integram a sociedade. Sendo assim, conforme descata Habermas, a teoria do véu da ignorância se baseia na assimetria informacional dos responsáveis pela decisão, eis que “as informações que se acrescentam a partir daí precisam estar em harmonia com os princípios já selecionados durante a subtração de informações, não pode haver maus imprevistos”.⁴⁰⁸ Dessa forma, “o ônus do processamento da informação, da qual o teórico exige as partes em sua condição primitiva, volta a incidir sobre ele mesmo”⁴⁰⁹, pois os conceitos básicos usados por Rawls na construção por meio do véu da ignorância (cidadão politicamente autônomo, cooperação justa e honesta e sociedade bem ordenada) não resistem à “revisão em face das experiências e dos aprendizados futuros”⁴¹⁰, o que mina a ideia defendida pela primeira fase da teoria rawlsiana. Logo, a concepção de objetividade advinda do duplo canal de correção (vide subcapítulo 1.6), que tem por base um procedimento de correção relativa⁴¹¹ (Alexy) e o entendimento racional de falantes na arena deliberativa (Habermas), contribui para a maior explicabilidade das decisões judiciais auxiliadas por sistemas de Inteligência Artificial, uma vez que se baseia no discurso do mundo da vida, ou seja, naquilo que realmente acontece e que permeia a construção da justificação racional da decisão judicial.

Para além disso, tendo por base a concepção de objetividade inerente ao duplo canal de correção, numa abordagem que olha para as dificuldades deste modelo de objetividade na implementação de julgamentos, destaca-se o estudo de Daniel

⁴⁰⁷ RAWLS, John. **A Theory of Justice**, p. 195-257.

⁴⁰⁸ HABERMAS, Jürgen. **A Inclusão do Outro**, p. 72.

⁴⁰⁹ Id.

⁴¹⁰ Id.

⁴¹¹ “O problema da correção leva, com isso, a um fendimento do conceito de correção em um conceito de correção procedimental absoluto e um relativo. Se tanto N como $\neg N$ são resultado de um procedimento, então ambos são, relativamente a este procedimento, corretos. O conceito procedimental absoluto da correção exige, pelo contrário, que, mais além, seja procurado por somente uma resposta. O conceito de correção procedimental relativa desempenha em discursos reais um papel central” (ALEXY, Robert. **Direito, Razão, Discurso**, p. 97).

Kahneman a respeito dos vieses⁴¹² e dos ruídos⁴¹³ cognitivos nos julgamentos humanos, inclusive nos julgamentos judiciais. Nessa linha, uma insuficiente percepção da realidade, causada pelas condições e características naturais ao pensamento humano, pode levar a decisões erradas, ainda que essas sejam revestidas por uma máscara de argumentos de justificação racional.

Desse modo, para além de onde os códigos legais alcançam, no âmbito da Teoria do Direito e da Filosofia do Direito há o debate sobre quais seriam os pressupostos para o julgamento adequado levando-se em conta o maior ou menor grau de objetividade do intérprete do Direito.

Isso ocorre porque a objetividade, sozinha, não garante correção na aplicação do Direito. Juízes erram, e o erro fica muitas vezes encoberto pelo argumento do convencimento motivado inerente à independência judicial. Nem mesmo as regras do art. 489, *caput*, II, e §1º do Código de Processo Civil são total garantia de racionalidade, pelo contrário são o mínimo que se espera de uma fundamentação analítica para a decisão judicial.

O direito processual tem seus limites e seu escopo, mas para a análise da decisão judicial auxiliada pela Inteligência Artificial um passo à frente deve ser dado neste estudo, de maneira a abordar os objetivos de sistemas de Inteligência Artificial aplicados no apoio à decisão judicial, a fim de que um mínimo de objetividade seja garantido.

A esse respeito, a noção de que a figura do juiz, por si só, garantiria um julgamento justo é bastante limitada e se esquece das características do Direito e da realidade fática (mundo ontológico).

Nos casos concretos, os Juízes, embora imbuídos de boa-fé, podem vir a errar em seus julgamentos, seja mediante erro na aplicação do Direito ou erro na análise probatória causados por vieses de julgamento que levam à inadequada compreensão

⁴¹² “A principal função do Sistema 1 é manter e atualizar um modelo de seu mundo pessoal, que representa o que há de normal nele. O modelo é construído por associações que ligam ideias de circunstâncias, eventos, ações e resultados que coocorrem com alguma regularidade, seja ao mesmo tempo, seja dentro de um intervalo relativamente curto” (KAHNEMAN, Daniel. Op. cit., posição 1498).

⁴¹³ KAHNEMAN, Daniel; SIBONY, Olivier; SUNSTEIN, Cass R. **Noise: A Flaw in Human Judgment**. New York: Little, Brown Spark, 2021, p. 26-28. E-book Kindle.

do caso⁴¹⁴, por ruídos surgidos⁴¹⁵ no processo e por viés cognitivo.⁴¹⁶ Por exemplo, chama a atenção o estudo realizado em Israel que aponta que a probabilidade de concessão de liberdade condicional varia de 75% no início da manhã para quase zero quando se aproxima do horário do almoço dos juízes.⁴¹⁷ Igualmente é preocupante a variabilidade de análises para casos julgados em condições análogas, pois conforme estudo realizado com Juízes Federais nos EUA, para o mesmo tipo de crime, cuja pena média era de 7 anos de restrição de liberdade, o desvio padrão na pena foi de 3,4 anos, o que “deve dar uma ideia da loteria que o réu enfrenta em um tribunal federal” norte-americano (tradução livre).⁴¹⁸

Nesse sentido, a percepção dos vieses e ruídos se enquadra na concepção de racionalidade limitada (*“bounded rationality”*), pois é preferível uma decisão satisfatória que diminua os erros de julgamento através de uma correção procedimental relativa⁴¹⁹, ao invés de uma decisão ótima que seja dificilmente alcançável⁴²⁰, uma vez que o máximo grau de racionalidade nem sempre se manifesta

⁴¹⁴ “A fé exagerada dos pesquisadores no que pode ser aprendido de umas poucas observações relaciona-se estreitamente com o efeito halo, a sensação que muitas vezes temos de conhecer e compreender uma pessoa sobre a qual na verdade sabemos muito pouco. O Sistema 1 se antecipa aos fatos ao construir uma imagem rica com base em fragmentos de evidência. Uma máquina de tirar conclusões precipitadas agiria como se acreditasse na lei dos pequenos números. De modo mais geral, vai produzir uma representação da realidade que faz sentido demais” (KAHNEMAN, Daniel. Op. cit., posição 2334).

⁴¹⁵ “Unfortunately, the world of federal justice is not perfect. The judges are not identical, and variability within columns is large, indicating noise in the judgments of each case. There is more variability in sentences than there should be, and the study’s aim is to analyze it” (KAHNEMAN, Daniel; SIBONY, Olivier; SUNSTEIN, Cass R. Op. cit., p. 70).

⁴¹⁶ Por exemplo: “Juizes alemães com uma média de mais de 15 anos de experiência em tribunal primeiro liam a descrição de uma mulher que fora detida por furto em lojas, depois lançavam dois dados que haviam sido adulterados de modo a dar sempre 3 ou 9. Assim que os dados paravam de se mover, perguntava-se aos juízes se iriam sentenciar a mulher a uma pena de prisão maior ou menor, em meses, do que o número apresentado no dado. Finalmente, os juízes eram instruídos a especificar a exata sentença de prisão que dariam à mulher. Em média, os que haviam rolado um 9 diziam que iriam sentenciá-la a oito meses; os que obtinham um 3 diziam que iriam sentenciá-la a cinco meses; o efeito de ancoragem foi de 50%” (KAHNEMAN, Daniel. Op. cit., posição 2571).

⁴¹⁷ “We have presented evidence suggesting that when judges make repeated rulings, they show an increased tendency to rule in favor of the status quo. This tendency can be overcome by taking a break to eat a meal, consistent with previous research demonstrating the effects of a short rest, positive mood, and glucose on mental resource replenishment (11–13). However, we cannot unequivocally determine whether simply resting or eating restores the judges’ mental resources because each of the breaks was taken for the purpose of eating a meal” (DANZIGER, Shai; LEVAV, Jonathan; Avnaim-Pesso, Liora. Extraneous factors in judicial decisions. **PNAS**, v. 108, n. 17, April 26, 2011. DOI: 10.1073/pnas.1018033108).

⁴¹⁸ KAHNEMAN, Daniel; SIBONY, Olivier; SUNSTEIN, Cass R. Op. cit. p. 71.

⁴¹⁹ ALEXY, Robert. **Direito, Razão, Discurso**, p. 99-100.

⁴²⁰ “When the course of action ‘than which none better’ is not defined, one will have to go for something that is good enough or satisfactory, rather than optimal. In the case of planning, this might involve substituting an ‘agreeable plan’ for the optimal plan” (ELSTER, Jon. **Sour Grapes: Studies in the Subversion of Rationality**. Cambridge: Cambridge University Press, 1983, p. 14. E-book Kindle).

na opção ótima.⁴²¹ Assim, o desenvolvimento de sistemas de Inteligência Artificial voltados ao julgamento pode se inspirar na teoria ampla da racionalidade de Jon Elster (*“broader theory of rationality”*)⁴²², pois a otimização da racionalidade ocorre de maneira gradativa dentro dos seus limites⁴²³, motivo pelo qual o conformismo⁴²⁴ com os limites da racionalidade se manifesta como uma forma de restringir a autonomia. Desse modo, o uso de Inteligência Artificial pode melhorar os julgamentos, mas não levará necessariamente à perfeição.

Logo, a autonomia é limitada pela realidade, o que limita a autonomia do próprio julgador. Por esse motivo, a IA voltada ao auxílio às decisões judiciais não é uma limitação à autonomia dos juízes, pois a autonomia dos julgamentos já é autolimitada pelas condições postas pelo mundo. Em suma, após a eliminação de desejos irracionais, a autonomia é um resíduo⁴²⁵. Dessa maneira, ao diminuir os vieses e ruídos de julgamento, a Inteligência Artificial deve ser implementada de maneira a preservar a autonomia residual dos julgadores humanos.

Assim, em sua relação com duplo canal de correção, a teoria da racionalidade limitada como fundamento da utilização de Inteligência Artificial nos julgamentos tem um apelo à humildade do julgador perante o mundo, pois a correção absoluta é dificilmente alcançável⁴²⁶, bem como leva em consideração que a simples

⁴²¹ “The fact that we have no reason for choosing between ‘maximum’ and ‘maximax’ does not imply that they are both optimal, as in the first case envisaged in the last paragraph. To assimilate the two cases would be to confuse indifference and incomparability” (Ibid., p. 13).

⁴²² “We need a broader theory of rationality that goes beyond the exclusively formal considerations of the preceding section, and that allows a scrutiny of the substantive nature of the desires and beliefs involved in action. We want to be able to say that acting rationally means acting consistently on beliefs and desires that are not only consistent, but also rational” (Ibid., p. 15).

⁴²³ “This means that when we leave the thin theory of rationality, the link between rationality and optimality is completely broken. Rational behaviour can be characterized as optimizing only – or at most – with respect to given beliefs about the world, but the principles of rationality governing belief acquisition cannot be spelled out in terms of optimization” (Ibid., p. 17-18).

⁴²⁴ “Autonomous moral action, indeed, implies the capacity to act morally even in non-moral environments. By this I do not mean that in such cases one should invariably follow the course that would have been best if adopted by all. By assumption this would not be conformist behaviour, yet it might be non-autonomous in some other way. Practical morality is largely about ‘second-best’ choices, meaning that when others act non-morally, one may have to deviate not only from what they do, but from the behaviour to be adopted in the ideal situation of universally moral behavior” (Ibid., p. 23-24).

⁴²⁵ “In the present work, autonomy will have to be understood as a mere residual, as what is left after we have eliminated the desires that have been shaped by one of the mechanisms on the short list for irrational preference-formation. Similarly, the quality of judgment for beliefs is also to be understood as the absence of distortions and illusions. This provides some necessary conditions for rationality in the broad sense which, while far less satisfactory than a full characterization, can at least be seen as a first step to that goal” (Ibid., p. 24).

⁴²⁶ ALEXY, Robert. **Direito, Razão, Discurso**, p. 100.

consistência é insuficiente.⁴²⁷ Nessa linha, a teoria da racionalidade limitada se manifesta através do que Elster chama de racionalização⁴²⁸, o que consiste em uma busca gradativa pela melhor decisão possível dentro do enquadramento que a realidade oferece ao julgador, levando-o a se abster de uma busca pelo resultado perfeito⁴²⁹, e aceitando os resultados possíveis com as condições que possui (e aqui a IA é uma aliada voltada à melhoria das condições de julgamento). Isso remete à alegoria da raposa que, ao não alcançar as uvas de um galho mais alto, imagina que as frutas são azedas e se conforma com a situação.⁴³⁰

A partir disso verifica-se que o erro de um juiz é tratado como um risco inerente ao processo judicial, pois inexistente Juiz imune a erros, salvo o mitológico juiz Hércules didaticamente idealizado por Ronald Dworkin, que tem a missão de construir um esquema de princípios, abstratos e concretos, que possibilite a “justificação coerente de todos os precedentes do direito costumeiro e, na medida em que estes devem ser justificados por princípios, também um esquema que justifique as disposições constitucionais e legislativas”.⁴³¹ Ou seja, o juiz que nunca erra e sempre encontra a resposta correta “fazendo valer suas aptidões extraordinárias”⁴³² existe apenas no plano das ideias. Ele seria equipado com habilidades sobre-humanas de sabedoria, paciência e acuidade, sendo que, na teoria de Dworkin, o “juiz real teria a tarefa de se aproximar a este ideal o mais possível”.⁴³³ A respeito do “juiz mítico” de Dworkin,

⁴²⁷ “I suggest that between the thin theory of the rational and the full theory of the true and the good there is room and need for a broad theory of the rational. To say that truth is necessary for rational beliefs clearly is to require too much; to say that consistency is sufficient, to demand too little. Similarly, although more controversially, for rational desires: the requirement of consistency is too weak, that of ethical goodness too strong” (ELSTER, Jon. **Sour Grapes**, p. 15).

⁴²⁸ “Rationalization. Sour grapes is a mechanism for dissonance reduction that operates on the preferences by which options are graded. An alternative mechanism works on the cognitive elements, shaping the perception rather than the evaluation of the situation. In some cases adaptive preferences and adaptive perception (i.e. rationalization) can hardly be distinguished from each other” (Ibid., p. 124).

⁴²⁹ “Similarly for counteradaptive preferences, where to ‘Forbidden fruit is sweet’ there corresponds ‘The grass is always greener on the other side of the fence.’ But in many cases the phenomena are clearly distinct. If I do not get the promotion I have coveted, I may ease the tension by saying that ‘my superiors fear my ability’ (misperception of the situation) or ‘the top job is not worth having anyway’ (misinformation of preferences). Or again I may change my life-style so as to benefit from the leisure permitted by the less prestigious position (character planning)” (Id.).

⁴³⁰ “In the French version of the sour grapes, cited in the epigraph to this chapter, the fox is deluded in his perception of the vermillion grapes, wrongly believing them to be green. In the English version he wrongly believes them to be sour, a matter of taste rather than of belief” (Id.).

⁴³¹ DWORKIN, Ronald. **Levando os Direitos a Sério**, p. 182.

⁴³² Ibid., p. 183.

⁴³³ “Un juez ideal, al que Dworkin llama ‘Hércules’, equipado con ‘superman skill, learning, patient and acumen’, es decir, con habilidad, sabiduría, paciencia e agudeza sobrehumanas, estaría en situación de encontrar la única respuesta correcta. Al juez real le corresponde la tarea de aproximarse a este ideal lo más posible” (ALEXY, Robert. *Sistema Jurídico, Principios Jurídicos y Razón Práctica*. **Doxa**, Madrid, n. 5, p. 140, 1988).

Frank Michelman destaca que ele é um solitário que não conversa com ninguém e que é heroico demais (o que demonstra falta de observância à pluralidade), nada abala o juiz Hércules, pois suas construções narrativas são monólogos.⁴³⁴ Por sua vez, Habermas rejeita parcialmente a ideia do juiz Hércules, haja vista que os juízes do mundo real são de carne e osso e acabam amalgamados pelos próprios preconceitos, que fazem representar em suas concepções de direito, não conseguindo se aproximar do modelo de reconstrução racional do Direito idealizado por Dworkin.⁴³⁵

A ideia de um juiz que não erra é uma fábula para qualquer operador do Direito que tenha experiência suficiente na realidade prática. Aqui se assume como erro não apenas equívocos na análise probatória, como também o erro na aplicação do direito.⁴³⁶ O que resta é a objetividade do próprio direito otimizada pelo caráter institucional de sua aplicação, necessidade apontada por autores como Merkl⁴³⁷,

⁴³⁴ *"This brings us, finally, to what is lacking in Ronald Dworkin's conception of law as (Judicial) integrity, even on my most optimistic reconstruction. What is lacking is dialogue. Hercules, Dworkin's mythic judge, is a loner. He is much too heroic. His narrative constructions are monologues. He converses with no one, except through books. He has no encounters. He meets no otherness. Nothing shakes him up. No interlocutor violates the inevitable insularity of his experience and outlook. Hercules is just a man, after all. He is not the whole community. No one man or woman could be that.*

Dworkin has produced an apotheosis of appellate judging without attention to what seems the most universal and striking institutional characteristic of the appellate bench, its plurality. We ought to consider what that plurality is "for." My suggestion is that it is for dialogue, in support of judicial practical reason, as an aspect of judicial self-government, in the interest of our freedom. There is a message there for the politics of judicial appointments, not to mention for the politics of law" (MICHELMAN, Frank. Foreword: Traces of Self-Government. **HARVARD LAW REVIEW**, Cambridge, v. 100, n. 1, p. 4-77, nov. 1986).

⁴³⁵ "Uma vez que os juízes, que são seres de carne e osso, ficam aquém da figura ideal de um Hércules, a recomendação de orientar-se por eles no trabalho diário nada mais é, na verdade, do que o reflexo do desejo de uma confirmação prática de decisão, que sempre é determinada através de interesses, enfoque político, parcialidade ideológica ou outros fatores externos. Juízes escolhem princípios e objetivos, construindo a partir daí teorias jurídicas próprias, a fim de 'racionalizar' decisões, isto é, dissimular preconceitos com os quais eles compensam a indeterminação do direito. Dworkin poderia responder a isso, explicando uma premissa deixada mais ou menos no escuro. Na medida em que os críticos podem realmente provar, apoiando-se em estudos de caso convincentes, que as decisões judiciais podem ser explicadas melhor a partir de fatores extraleais do que a partir da situação jurídica, os fatos falam contra a prática existente. Entretanto, a indeterminação interna do direito não resulta na estrutura do próprio direito, como os críticos pensam, e sim da recusa dos juízes em desenvolver a melhor teoria possível, bem como da história institucional de uma ordem jurídica, a qual é avessa, em maior ou menor medida, a uma reconstrução racional" (HABERMAS, Jürgen. **Direito e Democracia**, p. 266).

⁴³⁶ Erro na aplicação do direito, levando-se em conta o modelo jurídico adotado no momento da decisão.

⁴³⁷ "[No que diz respeito à] existência, o direito objetivo é *causado*, tal como tudo o que existe; [ele] é reconduzível a um sujeito. Mas é característico da interpretação subjetiva somente recorrer a esse sujeito, atribuir a ele *mesmo*, e não meramente a sua *expressão* um significado, tal como se recorre [a sua expressão] no caso de *dúvida*, quando não em *todo caso*. A lei é para ela [para a interpretação subjetiva] somente a manifestação, a ela [a lei] permanece para ela [para a interpretação subjetiva] sempre uma mera emanção. Para o intérprete subjetivo, é necessária e exclusivamente a instância competente [o que constitui] o princípio emanador" (MERKL, Adolf Julius. **Escritos de Teoria do Direito**. Organização e Tradução: Matheus Pelegriano da Silva. São Leopoldo: Editora UNISINOS, 2018, p. 136-137).

Kelsen⁴³⁸, Hart⁴³⁹ e MacCormick⁴⁴⁰, por exemplo.

Levando em conta o recorte da presente pesquisa, passa-se a tratar da objetividade do Direito como uma característica que estimula a existência dos sistemas de Inteligência Artificial judicial. Nesse sentido, pode-se perceber a objetividade possível no pensamento positivista⁴⁴¹, a busca pela objetividade limitada⁴⁴² que surge da correção procedimental relativa⁴⁴³ decorrente da

⁴³⁸ “A norma geral que, a certas condições determinadas, de modo abstrato tem de ser individualizada e concretizada para entrar em contato com a vida social, para ser aplicada à realidade. Para esse fim, tem-se de, num caso concreto, averiguar-se as condições, determinadas *in abstracto* na norma geral, possa ser ordenada e executada *in concreto*. Esses são dois elementos essenciais da função judicial. Essa função não tem, de modo algum, como às vezes se supõe, um caráter puramente declaratório. Contrariamente ao que às vezes se afirma, o tribunal não formula apenas um Direito já existente. Ele não ‘busca’ e ‘acha’ apenas o Direito que existe antes da decisão, não pronuncia meramente o Direito que existe, pronto e acabado, antes do pronunciamento”. [...] “A norma individual da decisão judicial é a individualização e a concretização necessárias da norma geral e abstrata” (KELSEN, Hans. **Teoria Geral do Direito e do Estado**. Tradução: Luís Carlos Borges. 4ª ed. São Paulo: Martins Fontes, 2005, p. 196-197).

⁴³⁹ As “formas particularizadas de controle ou são excepcionais, ou constituem acompanhamentos ou reforços subsidiários de formas gerais de instruções que não designam indivíduos específicos nem se dirigem a estes, e não especificam um ato em particular a ser praticado. Por isso, até a forma-padrão de uma lei penal escrita (que, entre todas as variedades do direito, é a mais semelhante a uma ordem apoiada por ameaças) é geral em dois sentidos: indica um tipo geral de conduta e se aplica a uma classe geral de pessoas, das quais se espera que estejam conscientes de que se aplica a elas e que a obedecem”. (...)

“Portanto, o controle jurídico consiste principalmente, embora não exclusivamente, no controle exercido por meio de instruções que são gerais nesse duplo sentido”. (...) “Num Estado moderno compreende-se normalmente que, na ausência de indicações especiais que ampliem ou restrinjam a classe, suas leis gerais se estendem a todas as pessoas no interior de seus limites territoriais” (HART, Herbert A. L. Op. cit., p. 26-27).

⁴⁴⁰ “Os juízes precisam fazer ‘justiça de acordo com a lei’, não justiça pura e simples. As normas do sistema jurídico fornecem uma concepção sólida da justiça que em circunstâncias normais – em que a justificação por dedução basta por si mesma – é cumprida em termos suficientes pela aplicação de normas pertinentes e aplicáveis segundo seus termos” (MACCORMICK, Neil. **Argumentação Jurídica e Teoria do Direito**. Tradução: Waldéa Barcellos. 2ª ed. São Paulo: WMF Martins Fontes, 2009, p. 93).

⁴⁴¹ “Se queremos caracterizar não apenas a interpretação da lei pelos tribunais ou pelas autoridades administrativas, mas, de modo inteiramente geral, a interpretação jurídica realizada pelos órgãos aplicadores do Direito, devemos dizer: na aplicação do Direito por um órgão jurídico, a interpretação cognoscitiva (obtida por uma operação de conhecimento) do Direito a aplicar combina-se com um ato de vontade em que o aplicador do Direito efetua uma escolha entre as possibilidades reveladas através daquela mesma interpretação cognoscitiva. Com este ato, ou é produzida uma norma de escalão inferior, ou é executado um ato de coerção estatuído na norma jurídica aplicanda” (KELSEN, Hans. **Teoria Pura do Direito**. Tradução: João Baptista Machado. 8ª ed. São Paulo: WMF Martins Fontes, 2009, p. 394).

⁴⁴² ELSTER, Jon. **Sour Grapes**, p. 13-24.

⁴⁴³ ALEXY, Robert. **Direito, Razão, Discurso**, p. 99-100.

racionalidade⁴⁴⁴ e o modelo da objetividade possível da integridade⁴⁴⁵ inerente à decisão de um juiz que se enxerga como agente constitucional apenas em questões de princípio (guardando deferência ao legislador em relação à escolha “fundamentalmente política”).⁴⁴⁶

Dessa maneira, a objetividade (em variados graus e com diversos matizes da Ciência Jurídica) é uma característica que se espera de qualquer julgador que pretenda possibilitar explicabilidade às suas decisões. Todavia a prática jurídica mostra que os julgamentos variam de maneira bastante evidente conforme variam os julgadores. Há nisso uma entropia do Direito na sua realização em cada julgamento (vide subcapítulo 1.5), pois no plano concreto, dentro de um espaço legítimo de decisão, há uma variabilidade na manifestação do jurídico que é intrínseca ao processo de individualização da norma.

Tal entropia é mitigada pela força aglutinadora dos Tribunais (e a respectiva carga de argumentação⁴⁴⁷) ao buscarem, por vezes, manter um sistema jurídico estável, íntegro e coerente através da jurisprudência e dos precedentes (art. 926 do Código de Processo Civil).⁴⁴⁸ Isso ocorre porque os juízes, quando individualmente considerados, atuam conforme sua independência funcional e julgam conforme a

⁴⁴⁴ “Se, desse modo, é introduzido, como peça de união entre procedimento e correção, o pressuposto de uma capacidade de juízo, em princípio existente, suficiente, isto é, uma capacidade de distinguir fundamentos bons de ruins para declarações substantivas, então se pergunta, por que, no fundo, ainda se direciona para o procedimento e não logo para fundamentos bons ou fundamentações suficientes. O fundamento para isso é que, em todo o caso, em perguntas práticas, nas quais se trata, essencialmente, de interpretação de interesses e compensações de interesses, não existem fundamentos bons, em si existentes. O que é um bom fundamento pode mostrar-se primeiro no processo de revisão discursiva. Se se quer empregar os conceitos de subjetividade e objetividade, então se pode dizer que o resultado do procedimento do discurso nem é só subjetivo nem só objetivo. Ele é objetivo, à medida que ele pôde resistir a uma revisão discursiva, que teve lugar sobre a base de uma capacidade de juízo, em princípio existente, dos participantes. Desse modo, a teoria do discurso evita tanto as fraquezas de teorias morais subjetivistas ou relativistas como das objetivas” (Ibid., p. 95-96).

⁴⁴⁵ DWORKIN, Ronald. **O Império do Direito**. Tradução: Jeferson Luiz Camargo. 3ª ed. São Paulo: Martins Fontes, 2014, p. 263-269.

⁴⁴⁶ [O juiz Hércules] “também não é um ‘ativista’. Vai recursar-se a substituir seu julgamento por aquele do legislador quando acreditar que a questão em jogo é fundamentalmente política, e não de princípio, quando o argumento for sobre as melhores estratégias para satisfazer inteiramente o interesse coletivo por meio de metas, tais como a prosperidade, a erradicação da pobreza ou o correto equilíbrio entre a economia e a preservação” (Ibid., p. 474-475).

⁴⁴⁷ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 238.

⁴⁴⁸ Frente ao recorte do presente tópico, aqui não se entrará no intrincado debate sobre a natureza e os elementos um sistema jurídico, estudo que ocorre, por exemplo, em: Kelsen, Hans. **Teoria Pura do Direito**. 8ª ed, p. 215-308; Kelsen, Hans. **Teoria Geral do Direito e do Estado**, p. 161-207; Hart, Herbert L. A. Op. cit., p. 129-159; Raz, Joseph. **O Conceito de Sistema Jurídico**, 2012; Alexy, Robert. **Conceito e Validade do Direito**, p. 101-112; Alchourrón, Carlos E.; Bulygin, Eugenio. Op. cit.

teoria do Direito que lhes pareça a mais adequada, a mais correta ou mais justa. Isso leva a existência de julgamentos de casos análogos que podem receber, muitas vezes, soluções diferentes. Assim, a atuação fracionada e atomizada dos Juízes sempre irá existir em algum grau, mesmo dentro de um sistema jurídico que adote precedentes vinculantes como força de coesão.

Nesse aspecto, a explicabilidade do modelo procedimental de decisão ganha em objetividade, na medida em que, ao se pautar na pretensão de correção relativa, leva a um caminho argumentativo e racional para a construção da uma solução para cada caso, paradigma que pode receber o auxílio da Inteligência Artificial judicial de maneira mais natural, sem que haja a substituição do intérprete humano.

Dessa forma, completude e não contradição são pretensões dos sistemas jurídicos que quando não realizadas levam à inconsistência do sistema (Bulygin⁴⁴⁹), à pluralidade de respostas possíveis (Hart⁴⁵⁰ e Alexy⁴⁵¹), às lacunas verdadeiras (Kelsen⁴⁵²) ou à abertura para que juízes construam justiça constitucional (Dworkin⁴⁵³ e Alexy⁴⁵⁴). Assim, nos limites de tensão entre o direito normativo e a realidade prática, a completude e a não contradição terão como resquício apenas os elementos de coesão mais básicos do sistema.

Dessa maneira, a entropia inevitável é o desafio à explicabilidade, pois a desorganização do que é organizado é um caminho inerente à teoria da informação⁴⁵⁵,

⁴⁴⁹ “A system is inconsistent in a case C_i if and only if there are two or more different (incompatible) solutions in the line corresponding to C_i . A system inconsistent if and only if there is at least one case in which it is inconsistent. Otherwise it will be called consistent” (ALCHOURRÓN, Carlos E.; BULYGIN, Eugenio. Op. cit., p. 17).

⁴⁵⁰ HART, Herbert L. A. Op. cit., p. 166-167.

⁴⁵¹ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 17-18.

⁴⁵² KELSEN, Hans. **Teoria Pura do Direito**. 8ª ed., p. 276.

⁴⁵³ DWORKIN, Ronald. **O Império do Direito**. 3ª ed., p. 453-455.

⁴⁵⁴ “No âmbito da abertura do direito, tudo aquilo em que se apoia e/ou deve apoiar-se quem aplica o direito para satisfazer a pretensão de correção integra o direito. Assim, os princípios – ainda que não possam ser identificados como princípios jurídicos em virtude dos critérios de validade da constituição – e os outros argumentos normativos que fundamentam a decisão tornam-se componentes do direito. A cláusula ‘apoia-se e/ou deve apoiar-se’ exprime a combinação entre a dimensão real e a dimensão ideal da aplicação do direito. Integram o direito tanto aqueles argumentos com os quais quem o aplica sustenta facticamente suas decisões, mesmo quando estas não satisfazem a pretensão à correção, quanto aqueles nos quais as decisões devem apoiar-se para satisfazer a pretensão de correção. Com isso, torna-se possível uma crítica jurídica decisória a partir do ponto de vista do direito” (ALEXY, Robert. **Conceito e Validade do Direito**, p. 154-155).

⁴⁵⁵ “IE suggests that there is something even more elemental than life, namely being – that is, the existence and flourishing of all entities and their global environment – and something more fundamental than suffering, namely entropy. The latter is most emphatically not the physicists’ concept of thermodynamic entropy. Entropy here refers to any kind of destruction, corruption, pollution and depletion of informational objects (mind, not of information), that is, any form of impoverishment of being, including nothingness, to phrase it more metaphysically. IE then provides a common vocabulary to

pois os juízes vão, cada um a seu modo, embaralhar e rescrever aquilo que o legislador e o constituinte fizeram, uma vez que a entropia ocorre naturalmente quando o julgador funciona como uma espécie de decodificador⁴⁵⁶ da expressão jurídica. Isso pode ser desejado ou indesejado, a depender da escola de pensamento de teoria do Direito, mas é um fato que acontece e que deve ser levado em conta quando se pensa nas possibilidades e desafios da interpretação e da aplicação do Direito com auxílio de sistemas de Inteligência Artificial. Isso para uma teoria que pretenda enfrentar o problema de maneira conectada com o que efetivamente acontece.

Assim, não é caso de se pensar a explicabilidade de uma Inteligência Artificial judicial “positivista” ou “pós-positivista” (com todas as variantes naturais a este termo), mas sim de se pensar a teoria do Direito para a explicabilidade da decisão judicial auxiliada por Inteligência Artificial que leve em conta aquilo que acontece. A “Torre de Babel” das decisões não será resolvida totalmente pela aplicação de Inteligência Artificial, pois a pluralidade persistirá e as abordagens necessárias à explicabilidade devem levar isso em consideração.

A IA judicial, portanto, funciona a partir de uma concepção de aglutinação da linha decisória que confere mais estabilidade à interpretação do Direito, na medida em que se trata de uma “estabilidade relaxada”⁴⁵⁷, na qual mudanças de julgamento podem ser implementadas pelo julgador que se utiliza de sua independência com coerência e mediante justificação adequada inerente à pretensão de correção procedimental relativa. Assim, a IA judicial não se pauta em um congelamento da interpretação do Direito em busca de um padrão específico voltado ao estabelecimento uma única explicação, pois um Judiciário que atua com objetividade não é sinônimo de fossilização jurisprudencial. Todavia, ao se ter a sugestão de decisões por meio de sistema de Inteligência Artificial deve existir um ônus superatório

understand the whole realm of being through an informational level of abstraction (see below). IE holds that being/information has an intrinsic worthiness. It substantiates this position by recognising that any informational entity has a Spinozian right to persist in its own status, and a Constructionist right to flourish, i.e. to improve and enrich its existence and essence. As a consequence of such “rights”, IE evaluates the duty of any moral agent in terms of contribution to the growth of the infosphere (see section 2.5 and 2.6) and any process, action or event that negatively affects the whole infosphere – not just an informational entity – as an increase in its level of entropy and hence an instance of evil” (FLORIDI, Luciano. Information Ethics, Its Nature and Scope. **SIGCAS Computer and Society**, v. 36, n. 3, Set. 2006, p. 26).

⁴⁵⁶ SHANNON, Claude E. A Mathematical Theory of Communication.

⁴⁵⁷ A inspiração na “estabilidade relaxada” aqui referida remete à engenharia aeronáutica, todavia se trata de analogia.

da decisão proposta, inerente à justificação racional, contexto no qual se insere o dever de explicabilidade.

Dessa forma, a explicabilidade, voltada à Inteligência Artificial judicial, busca fomentar uma decisão judicial em um contexto amplo, o que passa a exigir uma dupla faceta: [a] deve existir explicabilidade das soluções sugeridas pelos sistemas de Inteligência Artificial judicial; e [b] deve haver a exigência de que decisões judiciais que se afastem de soluções sugeridas pela IA judicial sejam ostensivamente justificadas pelo julgador. Referida dinâmica computacional-argumentativa, inerente ao duplo canal de correção, não impede a paulatina adaptação do direito ao caminhar da realidade histórica, o que ocorre em um equilíbrio entre segurança jurídica e justiça.⁴⁵⁸

Nessa linha, é necessário que o sistema jurídico garanta o acesso aos Tribunais⁴⁵⁹, a fim de que seja realizada a unificação de jurisprudência e haja garantia do duplo grau de Jurisdição, o que tenderá a filtrar as soluções concorrentes que nascem na decisão de cada Juiz para que seja possibilitada a construção de precedentes que servirão para fomentar a explicabilidade da Inteligência Artificial de auxílio à decisão judicial.

Estabelecida esta premissa inerente ao respeito à previsibilidade nos limites delineados pelo sistema normativo, no Judiciário o Direito sofre uma pressão por desconstrução e reconstrução através dos julgamentos mais variados, um devir entrópico de todos os casos que são julgados (e com os dados destes). Tentar congelar esse rio caudaloso com leis, precedentes, súmulas, teses de repercussão geral e recursos repetitivos é uma tarefa difícil senão impossível. Não será a IA que irá superar a tensão entre a objetividade generalizante da previsão legislativa e a subjetividade tópica na interpretação do Direito. Por este motivo, para se pensar em graus de explicabilidade da IA judicial deve-se levar em conta as dificuldades trazidas pela própria entropia jurídica na busca de um grau de objetividade⁴⁶⁰ dissociado do

⁴⁵⁸ BARBOZA, Estefânia Maria de Queiroz. Op. cit., p. 206.

⁴⁵⁹ CAPPELLETTI, Mauro; GARTH, Bryant. **Acesso à Justiça**. Tradução: Ellen Gracie Northfleet. Porto Alegre: Sérgio Fabris, 1988, p. 75-81.

⁴⁶⁰ “Embora existam diferenças entre as distintas abordagens da objetividade consideradas aqui, a semelhança mais abrangente entre elas reside no reconhecimento compartilhado da necessidade de encontrar um fundamento imparcial” (SEN, Amartya. Op. cit., p. 1167).

aspecto pessoal⁴⁶¹, mediante a transcendência da posição pessoal para uma leitura do mundo que tenha o potencial de ser universalizável.⁴⁶²

Nesse sentido, a análise de qualquer situação voltada à explicabilidade de decisões auxiliadas por IA judicial envolve a independência de pensamento⁴⁶³, o que possibilita que iniquidades de julgamento sejam superadas pela alteridade observacional de perspectivas distanciadas e próximas necessárias ao exercício do pensamento⁴⁶⁴ (e esse é o desafio da decisão judicial auxiliadas pela IA), o que remete às diferenças entre observar e agir⁴⁶⁵, sem que isso represente uma superioridade implícita entre teoria e prática.⁴⁶⁶ Dessa forma, percebe-se que a objetividade voltada ao incremento da explicabilidade guarda necessária relação com os limites e as possibilidades de representação do Direito através de sistemas de Inteligência Artificial, o que marca um ponto de comunicação entre este tópico e o tópico desta obra que trata do referido tema conexo (subcapítulo 3.1).

Assim, a inter-relação entre objetividade, explicabilidade e representação do Direito na IA judicial aumenta as possibilidades do intérprete no que tange à posicionalidade de observação voltada à compreensão do mundo e às situações que

⁴⁶¹ “A busca de algum tipo de compreensão do mundo que seja independente da posição é fundamental para o esclarecimento ético que pode ser procurado em uma abordagem não relacional” (Ibid., posição 3436).

⁴⁶² “O essencial aí não é a compreensão posicional, mas algum tipo de compreensão transposicional. Assumir um ‘ponto de vista desde lugar nenhum’ seria, obviamente, a ideia adequada nesse contexto” (Ibid., posição 3438).

⁴⁶³ “Mesmo quando um ponto de vista independente da posição é o apropriado para uma avaliação epistemológica, ética ou política, a realidade da dependência posicional das observações pode ter de ser levada em conta para explicar a dificuldade de conseguir uma compreensão posicionalmente não tendenciosa. Manter perspectivas posicionais pode ter um importante papel de dificultar às pessoas a superação de pontos de vista posicionalmente limitados” (Ibid., posição 3446).

⁴⁶⁴ “O argumento aqui tem relação com a defesa da ‘imparcialidade aberta’, invocando ideias como o dispositivo metodológico do observador imparcial de Adam Smith, buscando perspectivas distanciadas e próximas” (Ibid., posição 3457).

⁴⁶⁵ “Na filosofia pitagórica, o filósofo é apontado, por meio de simbolismo alegórico, como aquele que participa da vida como espectador dos acontecimentos, e não como aquele que se coloca em posição ativa no desenvolvimento de diversas atividades. A própria raiz etimológica do termo *theoría* descende diretamente de *theastai*, o que transmite a ideia de superioridade da observação com relação à participação” (BITTAR, Eduardo Carlos Bianca; ALMEIDA, Guilherme de Assis de. **Curso de Filosofia do Direito**. 11ª ed. São Paulo: Atlas, 2015, p. 6).

⁴⁶⁶ “Coisa muita diversa ocorrerá numa época de crise. Sempre que o sistema ‘clássico’ do dia for ameaçado pelos resultados de experimentos novos, passíveis de serem interpretados como faseamentos, segundo o meu ponto de vista, o sistema permanecerá inabalado aos olhos do convencionalista. Ele afastará as incoerências que possam ter surgido, aludindo talvez, ao fato de não dominarmos suficientemente o sistema. Ou eliminará as incoerências, sugerindo a adoção de certas hipóteses auxiliares *ad hoc* ou, talvez, de certas correções nos instrumentos de medidas” (POPPER, Karl. Op. cit., p. 70-71).

o cercam⁴⁶⁷, o que contribui para o incremento informacional voltado à tomada de decisão. Nesse aspecto, a Inteligência Artificial judicial tem o potencial de se inserir como mais um elemento de objetividade para a análise dos casos. A partir desse modelo, a consciência humana e a objetividade possível⁴⁶⁸ da Inteligência Artificial atuam em cooperação, auxiliando-se mutuamente a fim de buscar ampliar as possibilidades de uma decisão correta.

Muitas vezes, esse incremento informacional fomentado pela necessidade de explicabilidade é alcançado a partir do desprendimento entre o ato de julgar e tudo repensar. Em um sistema de Inteligência Artificial judicial as ontologias⁴⁶⁹ e as heurísticas⁴⁷⁰ racionalmente estabelecidas têm seus papéis num julgamento adequado, ao servirem de mastro forte para que o julgador se amarre, a fim de resistir ao canto das sereias⁴⁷¹, de modo a evitar o uso abusivo da IA nos julgamentos mediante uma fuga do dever de explicabilidade através de uma máscara para os fundamentos que são escondidos. Nesse sentido, em respeito à exigência de saturação argumentativa⁴⁷², somente a justificação racional pautada na pretensão de correção pode superar as amarras trazidas em decisões sugeridas via Inteligência Artificial.

Assim, a escolha objetiva voltada à explicabilidade passa pela maximização da melhor razão que pode ser promovida no julgamento⁴⁷³, na busca de razões

⁴⁶⁷ “Embora a posicionalidade da observação e da construção desempenhe um papel importante no processo de avanço do conhecimento científico, ela é mais amplamente significativa na formação da opinião em geral: na compreensão social, bem como na busca das ciências naturais. Na verdade, o papel da posicionalidade pode ser particularmente importante na interpretação das ilusões persistentes e sistemáticas que podem influenciar — e distorcer — de forma significativa a compreensão social e a avaliação dos assuntos públicos” (SEN, Amartya. Op. cit., posição 3539).

⁴⁶⁸ Todavia, como será destacado no decorrer do texto, o cuidado é necessário, pois por trás de algumas “objetividades” podem estar escondidas subjetividades.

⁴⁶⁹ WALKER, Joshua. **On Legal AI: Um Rápido Tratado sobre a Inteligência Artificial no Direito**. Tradução: José Ignacio Coelho Mendes Neto. São Paulo: Revista dos Tribunais, 2021, p. 57-59.

⁴⁷⁰ RUSSELL, Stuart; NORVIG, Peter. Op. cit., p. 91-95.

⁴⁷¹ ELSTER, Jon. Ulysses and the sirens: A theory of imperfect rationality. **Social Science Information**, v. 16, n. 5, 469-526, 1977. DOI: 10.1177/053901847701600501.

⁴⁷² ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 216.

⁴⁷³ “A maximização é invocada na economia e nas ciências sociais principalmente como uma característica comportamental, mas é interessante notar que o ‘princípio do tempo mínimo’ de Fermat em óptica (tratando do modo mais rápido que a luz pode viajar de um ponto a outro), um refinado exercício de minimização, não constituía um caso de comportamento consciente, pois não há vontade alguma envolvida na ‘escolha’ da luz de viajar pelo caminho de tempo mínimo de um ponto a outro. Na verdade, na física e nas ciências naturais, a maximização normalmente ocorre sem um ‘maximizador’ voluntário. A ausência da escolha decisiva também se destina em geral às primeiras aplicações analíticas da maximização e da minimização, incluindo as geométricas, que remontam à busca do “arco mais curto” por matemáticos gregos, e a outros exercícios considerados pelos “grandes geômetras” no mundo antigo, como Apolônio de Pérgamo. Ao contrário, o processo de maximização na economia é visto principalmente como resultado de uma escolha consciente (mesmo que às vezes seja dado um

adequadas⁴⁷⁴ entre as que podem ser estimuladas por um sistema de Inteligência Artificial judicial e aquelas que são usadas pelo próprio julgador para manter a sugestão trazida pela IA ou para superá-la. Essa característica é um estímulo para o funcionamento do duplo canal de correção, pois possibilita decisões judiciais auxiliadas pela IA no paradigma do juiz ciborgue adotado nesta obra.

2.3 IA Judicial Explicável: Direito à Explicabilidade e Deliberação Racional

A explicabilidade dos resultados do processamento advindo de algoritmos de IA passa necessariamente pela ideia de *accountability*, sendo a explicabilidade um aprimoramento dessa abordagem.

Dessa maneira, a *accountability* no desenvolvimento de modelos de IA judicial deve passar pela transparência no processo construção do modelo algorítmico⁴⁷⁵, pois a explicabilidade advinda da transparência no procedimento de funcionamento da IA é a abordagem mais efetiva que se apresenta⁴⁷⁶, e pode ser dividida em fases de

papel ao ‘comportamento maximizador habitual’), e o exercício da escolha racional é geralmente interpretado como a deliberada maximização daquilo que uma pessoa tem a melhor razão para promover. Como Jon Elster expressa em seu breve, conciso e elegante livro *Reason and rationality*, “o agente racional é aquele que age por razões suficientes”. É realmente difícil evitar o pensamento de que a racionalidade da escolha deve ter uma forte ligação com o raciocínio. E é por causa da crença, frequentemente implícita e não explícita, de que o raciocínio tende a favorecer a maximização do que queremos promover ou buscar (uma ideia que não tem nada de extravagante) que a maximização é considerada fundamental para o comportamento racional” (SEN, Amartya. Op. cit. p. 3715-3732).

⁴⁷⁴ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 261-264.

⁴⁷⁵ “Our definition of transparency arguably allows for a wide range of combinations of A, B, C, and M. Inspired by Mansbridge (2009), we argue that in relation to transparency in public decision-making, a distinction can be made between transparency that (1) informs C (e.g., the public) about final decisions or policies; (2) about the process resulting in the decisions (transparency in process); and (3) about the reasons on which the decision is based (transparency in rationale). These forms of transparency should be understood as degrees rather than separate elements, as it is difficult to provide the reasons for a decision without making explicit what the decision is, in the same way that it is difficult to present the process leading up to a decision without making the reasons on which it is based explicit.

Thus, in most cases, the transparency of a process should be considered more transparent than the transparency of the reason, which, in turn, is more transparent than the transparency of the decision” (LICHT, K. de Fine; LICHT, J. de Fine. Artificial intelligence, transparency, and public decision-making. **AI & Society**, Mar. 2020. DOI: 10.1007/s00146-w).

⁴⁷⁶ “We believe that if perceived legitimacy is the goal, we should opt for transparency in rationale and not transparency in process. By transparency in rationale, we refer to the public receiving information for the justification or explanation of a decision as well as details on who can be held accountable for said decision. Thus, our meaning of transparency in rationale is similar to that of Floridi et al. (2018: 699f) and their use of “explicability,” which implies that the public receives an explanation or justification for the decision made, a description of the process leading up to it, and an account of who is responsible for it. However, if explicability means that we actually make the decision making processes fully transparent, then we do not believe it suitable in relation to the production of perceived legitimacy; but, if decision-makers should provide an explanation in the form of a narrative where it is explained how the decision has been made, then this might be applicable for perceived legitimacy” (Id.).

garantia da explicabilidade de modelos de IA⁴⁷⁷, o que consiste na definição do objetivo (primeira fase)⁴⁷⁸, na escrita do código algorítmico (*coding* – segunda fase)⁴⁷⁹ e na implementação do modelo (terceira fase).⁴⁸⁰ Dessa forma, pode-se compreender que a explicabilidade do processo de construção de modelos de IA tem pressupostos característicos que são distintos da explicabilidade das saídas (*outputs*) e dos resultados (*outcomes*) propiciados por um modelo de IA.

Nessa linha, no processo de construção das possibilidades de explicabilidade dos modelos de Inteligência Artificial judicial deve existir deliberação consciente acerca dos objetivos e das prioridades adotadas pelos desenvolvedores, parâmetros que devem ser avaliáveis na esfera pública.⁴⁸¹ Para além disso, no momento em que a IA é implementada, o código-fonte e os registros de como a IA é utilizada necessitam ser disponibilizados publicamente, conforme foi regulado pelo Conselho Nacional de Justiça mediante o depósito dos modelos de IA judicial no *Sinapses* (art. 10, III, do art. 332/2020 do CNJ), o qual é solução computacional para “armazenar, testar, treinar, distribuir e auditar modelos de Inteligência Artificial” (art. 3º, III, do art. 332/2020 do CNJ).⁴⁸²

⁴⁷⁷ “To appreciate what full transparency in AI and public decision-making would amount to, we propose dividing the entire decision-making process into three phases: Phase 1 is the goal-setting phase (goal-setting), Phase 2 is the coding phase (coding), and Phase 3 is the implementation phase (implementation) (see, e.g., de Laat: 529–533 for a comparison of AI in the market sphere and Boscoe 2019 for a similarly structured process in public decision-making)” (Id.).

⁴⁷⁸ “During Phase 1 (the goal-setting phase), decision-makers decide on the goals of the AI, how they should be weighed against each other when in conflict, and the features and data available to draw inferences from.” (...) “Of course, these features might pull in different directions, meaning that you will have to give the values different weights to guide the AI on what to do when said features are in conflict. These decisions are often highly political, as they require decision-makers to explicitly distinguish between advantages and disadvantages at a high level of precision” (Id.).

⁴⁷⁹ “In Phase 2 (the coding phase), the AI is developed and worked on to ensure it meets the necessary standards. This is often a point of introduction for problems related to bugs and biases, and is well described in the literature (e.g., Sweeney 2013; Datta et al. 2015; O’Neil 2016; Boscoe 2019). In this phase, it is discussed what the accuracy rates are, what they should be, how these and other performance metrics differ, how they should be allowed to differ across different subpopulations (when deciding about groups or individuals), what data to use when training the algorithm, and how to clean it” (Id.).

⁴⁸⁰ “In Phase 3 (the implementation phase), the AI is applied in the public decision-making processes, and the results produced by the AI are used in actual decision-making. This can be done by having the AI make the decision by itself or, more plausibly, by having an individual formally make the decision based on the results or recommendations of the AI. Naturally, this phase is often the one to which researchers refer when discussing AI and transparency. This phase will also, of course, feed back into Phases 1 and 2” (Id.).

⁴⁸¹ “To make the process of Phase 1 (goal-setting) and 2 (coding) fully transparent to the public (C), the decision-makers (A) need to make the deliberations about goals and prioritizations as well as the deliberations of the programmers (B) available to the public, along with the training data, the testing data etc” (Id.).

⁴⁸² CONSELHO NACIONAL DE JUSTIÇA. Resolução nº 332, de 21 de agosto de 2020.

Dessa maneira, a tomada de decisão judicial com o auxílio de IA pode ser mais célere e eficiente, sem que necessariamente haja perda da qualidade deliberativa, uma vez que a habilidade de chegar a uma decisão e implementá-la nos algoritmos é importante na tomada de decisões através de IA. Assim, desde que garantido o posterior controle e a supervisão por humanos (juizador e sua equipe técnica)⁴⁸³, eventuais vieses ou saídas incorretas (sem acurácia)⁴⁸⁴ podem ser tratados antes de ser efetivamente adotada a decisão sugerida pela IA judicial, devendo o erro da IA ser registrado no sistema (art. 25, Parágrafo único, V, da Res. 332/2020⁴⁸⁵) para que exista a transparência da decisão tomada.

Nesse ponto, a não realização do registro serviria como vulnerabilidade para a burla à proposta decisória advinda da IA sem a necessidade de demonstração dos motivos de tal superação, indicando, em quais pontos e por quais motivos, a decisão puramente humana do juizador (agora auxiliado apenas por sua equipe técnica) é juridicamente mais adequada que a proposta de decisão da IA que veio a ser descartada. Além disso, a IA judicial⁴⁸⁶ cumpre as funções de apontamento de vieses do juizador⁴⁸⁷ (o que é interesse da democracia) e de demonstração de casos em que ocorre a fuga injustificada da *ratio decidendi* de precedentes. Nesse sentido, a IA judicial cria um passo justificatório que deve ser publicamente acessível, a fim de que a tentação à simples desconsideração da racionalidade (ou irracionalidade) proposta pela IA seja inibida.

O grau de acesso às informações deve ser pensado de modo que as partes e os advogados tenham automaticamente informações sobre o que precisam saber acerca das principais variáveis que influenciam cada julgamento com IA⁴⁸⁸, uma vez

⁴⁸³ HARTMANN PEIXOTO, Fabiano; SILVA, Roberta Zumblick Martins da. Op. cit., p. 40.

⁴⁸⁴ “*Notwithstanding its potential benefits, the use of machine learning has prompted a number of concerns, especially when the systems make predictions that affect people’s liberty, safety, or privacy. One strand of criticism focuses on the ways in which these algorithms can replicate and exacerbate societal biases in light of the data on which scientists train them. Another line of critiques questions the accuracy of various machine learning predictions, with objectors claiming that tools such as criminal justice algorithms predict recidivism less accurately than humans*”. (DEEKS, Ashley. Op cit., p. 1833).

⁴⁸⁵ CONSELHO NACIONAL DE JUSTIÇA. Resolução nº 332, de 21 de agosto de 2020.

⁴⁸⁶ Seja um sistema/modelo de IA mais ou menos complexo, neste caso o foco é na sua aplicabilidade para o aprimoramento da decisão judicial, pois a IA judicial deve ser construída em benefício da sociedade da correta aplicação do Direito, não apenas para “acelerar” processos. O respeito à razoável duração do processo mediante a aceleração dos atos processuais é apenas uma das facetas positivas da implementação de sistemas de IA judicial. Otimização do processo decisório é mais, muito mais, do que apenas julgar rapidamente.

⁴⁸⁷ SUNSTEIN, Cass. Algorithms, Correcting Biases.

⁴⁸⁸ “*The second argument in favor of full transparency is that it should increase public understanding of decisions and decision-making processes, thus making the public more confident about decision-*

que caso haja a demanda por informações adicionais, estas sejam disponibilizadas para se garantir a legitimidade⁴⁸⁹ democrática inerente à transparência completa (*full transparency*), pois, mesmo que os destinatários da decisão não consigam compreender a forma de funcionamento de um sistema de IA judicial (um modelo baseado em rede neural, por exemplo), há a percepção de que o Judiciário, instituição democraticamente legítima⁴⁹⁰ para a função, respeita ativamente a transparência em seus julgamentos.⁴⁹¹ Em resumo, o Judiciário da era da IA não pode ter receio do escrutínio público, ainda que existam dificuldades de compreensão do funcionamento de modelos de Inteligência Artificial judicial e do próprio Direito.

Para além do cuidado em relação aos riscos, os quais devem ser catalogados, conhecidos e monitorados de maneira ativa, a IA de auxílio ao julgamento pode incorporar a extensão da legislação aplicável a determinada classe de problemas e, inclusive, a determinados valores para dentro do modelo⁴⁹², seja numa abordagem de *machine learning* (com treinamento em extensa base de dados) ou mediante representação do conhecimento nos próprios algoritmos, pautando-se sempre na possibilidade de pluralidade democrática (art. 7º da Resolução 332/2020 do CNJ⁴⁹³).

makers. There are several problems with this argument, but the most pertinent is what is sometimes called the information overload (e.g., Eppler and Mengis 2004) or the transparency paradox (Richards and King 2013)” (LICHT, K. de Fine; LICHT, J. de Fine. Op. cit.).

⁴⁸⁹ *“The third argument in favor of full transparency is that it increases perceived legitimacy because it induces a feeling of control among the public. The idea of a close relationship between transparency and accountability underlies many discussions of transparency (e.g., Gutmann and Thompson 1996; Hood 2010; Kosack and Fung 2014). In principal-agent terminology (Holmström 1979; Fearon 1998; Manin et al. 1999), transparency allows the principal (i.e., the public) to overcome the information asymmetry regarding the agent’s (i.e., the representative’s) workings, thereby leading to renewed instructions about what to do or even removal from office” (Id.).*

⁴⁹⁰ LORENZETTO, Bruno Meneses; CLÈVE, Clèmerson Merlin. Teorias Interpretativas, Capacidades Institucionais e Crítica.

⁴⁹¹ *“Hence, transparency in itself might have an independent positive effect on perceived legitimacy, irrespective of the content of the process that is being made available to the public. Thus, even if an individual cannot understand the functioning of a classifier (e.g., a deep neural network) or how the input generates the output, they could still appreciate the fact that the government has made the process transparent” (LICHT, K. de Fine; LICHT, J. de Fine. Op. cit.).*

⁴⁹² *“Another opportunity for the use of machine learning algorithms in the agency setting might be to parse and summarize voluminous public comments provided as part of notice and comment rulemaking. Further, as noted above, agencies might turn to machine learning to help them conduct adjudications. This could include using algorithms to predict pilot competence and grant pilot’s licenses, forecast the effects of a proposed merger on competition, or decide disability claims. None of these processes will exclude the human role entirely – at the very least, computer scientists must code agency ‘values’ into the algorithms in the form of ones and zeros - ut machine learning-driven rulemaking and adjudication may embody a host of decisional steps that are nontransparent and difficult to trace” (DEEKS, Ashley. Op. cit., p. 1833).*

⁴⁹³ CONSELHO NACIONAL DE JUSTIÇA. Resolução nº 332, de 21 de agosto de 2020.

Portanto, o projeto, o desenvolvimento e a implementação de modelos de IA judicial devem se basear no uso público da razão, inerentes à democracia e ao Direito, o que estimula uma concepção de explicabilidade que faz parte de um procedimento dialético-argumentativo e também traz consigo elementos materiais de um Judiciário democrático em seu processo decisório.

2.4 Deliberação Racional e IA Judicial: a Explicabilidade como Elemento da Justiça

A necessidade do desenvolvimento de sistemas Inteligência Artificial explicável (xAI) para os modelos aplicados ao Judiciário se amolda à necessária construção de decisões judiciais mediante a exposição da racionalidade discursiva⁴⁹⁴, sendo que esta tem o potencial de trazer ao centro do debate democrático aspectos, desde os mais básicos, que influenciam nos posicionamentos dos julgadores.

Nesse caminho, a ideia de explicabilidade, como elemento das razões compartilhadas nas decisões auxiliadas pela IA, é uma possibilidade que se adequa à democracia constitucional, pois, conforme destaca Habermas, a substância dos direitos está contida no procedimento argumentativo democrático “necessário à institucionalização jurídica do uso público da razão por parte dos cidadãos soberanos”.⁴⁹⁵ Este posicionamento garante a mais ampla explicabilidade democrática, uma vez que se apresenta mais inclusivo do que outras propostas filosóficas, como a presente na segunda fase de pensamento de John Rawls⁴⁹⁶, por exemplo.

A explicabilidade nasce como um elemento intrínseco ao uso público da razão, na medida em que o Judiciário, órgão constitucional de quem se espera

⁴⁹⁴ “O princípio do discurso submete a validade de qualquer tipo de norma de ação ao assentimento daqueles que, na qualidade de atingidos, tomam parte em ‘discursos racionais’. Na medida em que o princípio do discurso se aplica às normas de comportamento, que regulam interações simples num círculo ilimitado de destinatários, surgem questões que correspondem a um determinado tipo de discurso moral. Quando o princípio do discurso encontra aplicação em normas de ação que podem assumir a forma jurídica, entram em jogo questionamentos políticos de várias espécies. À lógica desses questionamentos correspondem diferentes tipos de discurso e formas de negociação” (HABERMAS, Jürgen. **Direito e Democracia**, p. 199).

⁴⁹⁵ HABERMAS, Jürgen. **A Inclusão do Outro**, p. 87.

⁴⁹⁶ [A] “razão pública é a razão de cidadãos iguais que, enquanto corpo coletivo, exercem um poder político final e coercitivo uns sobre os outros ao promulgar leis e emendar sua constituição” (RAWLS, John. **O Liberalismo Político**, p. 263).

racionalidade⁴⁹⁷ e sinceridade⁴⁹⁸ inerente à boa-fé argumentativa⁴⁹⁹, tem o ônus de explicar os motivos de suas decisões (art. 93, IX, e art. 489, *caput*, II, §1º e §2º, do CPC). Dessa maneira, devido aos aspectos sensíveis adicionais trazidos pela inclusão da IA judicial no processo decisório, por meio de racionalidade discursiva⁵⁰⁰ é necessário o estímulo à construção das decisões que possam ser razoavelmente aceitas na sociedade.⁵⁰¹

O processo de decisão através do uso público da razão leva à formação de consenso acerca das decisões judiciais (embora este não seja o fator determinante da correção⁵⁰², ainda que a influencie – vide o subcapítulo 1.6), consenso esse que será a fonte de inspiração para a programação do sistema de IA e/ou para os critérios de seleção e tratamento das bases de dados utilizadas no aprendizado de máquina. Tal consenso não é estático, mas sim dinâmico, uma vez que se trata de construção reiterada da racionalidade discursiva numa democracia viva, na qual há a construção

⁴⁹⁷ A “teoria do discurso oferece um critério, em situações específicas, para a racionalidade de processos de decisão e para a racionalidade das decisões produzidas neles” (ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 253).

⁴⁹⁸ “Um acordo sobre seu conteúdo mínimo se alcançaria rapidamente. Aqui se inclui, por exemplo, a exigência de consistência (não contradição), de racionalidade dos fins e de verdade dos enunciados empíricos empregados” (Ibid., p. 252).

⁴⁹⁹ PIRES, Teresinha Inês Teles. Op. cit.

⁵⁰⁰ “Por referir-se tanto à facticidade da validade social, que se mede pela obediência geral à normas, quanto à legitimidade da pretensão ao reconhecimento normativo, o direito permite aos membros da comunidade jurídica escolherem entre dois enfoques distintos em relação à mesma norma: objetivador e performativo, podendo adotar modos de ler correspondentes a cada um dos enfoques. Para o ‘arbitrio’ de um ator que se orienta pelo sucesso próprio, a regra constitui um empecilho fático na expectativa da imposição do mandamento jurídico – com consequências previsíveis, no caso de uma transgressão da norma. Quando, porém, se deseja entender-se com outros atores sobre condições a serem preenchidas em comum para que tenha sucesso em suas ações, a regra amarra a sua ‘vontade livre’ através de uma pretensão de validade deontológica. O fato de esta perspectiva ficar aberta não significa necessariamente uma fusão de momentos, pois, na perspectiva do ator, eles *continuam* irreconciliáveis. Ao variar a perspectiva escolhida, a norma jurídica forma um outro tipo de componente da situação: para o que age estrategicamente, ela se encontra no nível de fatos sociais que limitam extrenamente o seu espaço de opções; para o que age comunicativamente, porém, ela se situa no nível de expectativas obrigatórias de comportamento, em relação às quais se supõe um acordo racionalmente motivado entre parceiros jurídicos” (HABERMAS, Jürgen. **Direito e Democracia**, p. 51).

⁵⁰¹ “2. A democracia envolve uma relação política entre cidadãos no interior da estrutura básica da sociedade na qual nasceram e na qual normalmente passam toda a sua vida; isso implica ainda uma parte igual no poder político coercitivo que os cidadãos exercem uns sobre os outros ao votar, e de outras formas também.

Enquanto razoáveis e racionais, e sabendo-se que endossam uma grande diversidade de doutrinas religiosas e filosóficas razoáveis, os cidadãos devem estar dispostos a explicar a base de suas ações uns para os outros em termos que cada qual razoavelmente espere que outros possam aceitar, por serem coerentes com a liberdade e igualdade dos cidadãos. Procurar satisfazer essa condição é uma das tarefas que esse ideal de política democrática exige de nós” (RAWLS, John. **O Liberalismo Político**, p. 266-267).

⁵⁰² ALEXY, Robert. **Direito, Razão, Discurso**, p. 95.

das estruturas jurídico-organizacionais do Estado e da sociedade de maneira compartilhada.⁵⁰³

Nessa linha, mediante o uso público da razão deve existir a exposição consciente dos motivos das decisões tomadas com auxílio de Inteligência Artificial: **a)** pelos quais determinada solução (de representação de conhecimento ou de técnica de *machine learning*) foi adotada em detrimento de outras abordagens de IA; **b)** de por que determinada base de dados foi utilizada no treinamento, no teste ou na validação de modelo baseado em aprendizado de máquina; **c)** das causas de descontinuação ou não implementação de modelos de IA judicial; e, **d)** na medida da possibilidade da técnica, por que determinada decisão foi proposta por modelo de IA judicial (resultado do treinamento, espécie de estruturação algorítmica adotada, base de dados utilizada, escolha do juiz-desenvolvedor, escolha centralizada pela administração do Tribunal, etc.).

Desse modo, a abordagem através da construção democrática do uso público da razão tem o potencial de otimizar o direito à explicabilidade, animando-o com a exigência de um conteúdo efetivamente democrático ao respeitar direitos fundamentais, conteúdo esse que possa ser aceito como adequado à pretensão de validade⁵⁰⁴, no plano político-institucional (o que também leva em conta a pluralidade), e à pretensão de correção, no plano jurídico-argumentativo.⁵⁰⁵ Nesse sentido, o uso público da razão é uma garantia de que no processo decisão sobre a adoção da IA e nas decisões judiciais auxiliadas por IA haverá um ônus democrático que deverá ser cumprido pelos tribunais, em otimização à explicabilidade, o que tem o potencial de inibir abusos e a tentação de adoção de modelos de IA judicial voltados apenas para

⁵⁰³ “Com a força ilocucionária de uma externação, o falante pode motivar o ouvinte a aceitar sua oferta de um ato de fala e, com isso, motivá-lo a *estabelecer uma ligação racionalmente motivada*. Tal concepção prevê que sujeitos aptos a falar e agir possam fazer referência a mais que um único mundo; e que, ao se entenderem uns com os outros sobre alguma coisa em um mundo único, embasem sua comunicação sobre um sistema de mundos que suponham de maneira compartilhada. Nesse contexto, sugeri discernir o mundo exterior em dois, um mundo objetivo e um mundo social, e introduzir o mundo interior como conceito complementar a esse mundo exterior. As respectivas pretensões de validade de verdade, correção e veracidade podem então servir de fios condutores à escolha de pontos de vistas teóricos sob os quais possam fundamentar os modelos básicos de emprego da linguagem – as funções da linguagem – e classificar as ações de fala que variam de uma língua a outra” (HABERMAS, Jürgen. **Teoria do Agir Comunicativo**, 1, p. 484).

⁵⁰⁴ Ibid., p. 147-148.

⁵⁰⁵ “Quem afirma algo quer não só expressar que acredita em algo; ele alega algo além disso, a saber, que aquilo que ele fala também é fundamentável, verdadeiro ou correto. Isso vale da mesma forma para enunciados normativos e não-normativos” (ALEXY, Robert. **Teoria Discursiva do Direito**. 3ª ed. Tradução: Alexandre Travessoni Gomes Trivisonno. Rio de Janeiro: Forense, 2018, p. 48. E-book Kindle).

a produtividade, numa perspectiva numérica, em detrimento do desenvolvimento de sistemas computacionais que tragam mais segurança através de uma arquitetura de escolhas⁵⁰⁶ que estimule a decisão correta (racional)⁵⁰⁷, explicitamente justificada⁵⁰⁸ e consciente.

O referido processo público de construção da explicabilidade dos modelos de IA judicial envolve: **a)** a política de tecnologia da informação de cada Tribunal; **b)** o projeto do sistema de IA judicial; e **c)** a implementação de técnicas de explicabilidade (xAI) para os resultados (*outputs*) de processamento de decisões via Inteligência Artificial.

Ademais, em relação à explicabilidade das decisões auxiliadas por IA, o escrutínio no processo de desenvolvimento dos sistemas de IA judicial deve ser cauteloso, de maneira a não discriminar minorias ou a adotar viés sectário (art. 7º e art. 23, III, da Resolução 332/2020 do CNJ⁵⁰⁹), visto que o direito é um meio “para a auto-organização de comunidades jurídicas que se afirmam, num ambiente social, sob determinadas condições históricas”⁵¹⁰ (contexto no qual razões de determinados grupos atomizados podem ser necessárias para o estímulo do debate⁵¹¹), o que se faz

⁵⁰⁶ “Claro que é possível manipular a visibilidade, e bons arquitetos de escolhas podem direcionar a atenção das pessoas para os incentivos” (THALER, Richard H.; SUNSTEIN, Cass R. Op. cit., posição 1796).

⁵⁰⁷ “A explicação do conceito de argumentação jurídica racional se dá nesta investigação mediante a descrição de uma série de regras a serem seguidas e de formas que devem ser adotadas pela argumentação para satisfazer a pretensão que nela se formula. Se uma discussão corresponde a essas regras e formas, o resultado alcançado pode ser designado como “correto”. As regras e formas do discurso jurídico constituem por isso um critério de correção para as decisões jurídicas” (ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 253).

⁵⁰⁸ BRETAS, Carla Panza; MATTOS, Kennedy Josue Greca de; SCHIER, Paulo Ricardo. Teoria do Direito e da Constituição no contexto do neoconstitucionalismo pós-positivista: análise quanto à racionalidade. **Revista da Faculdade de Direito da UFRGS**, Porto Alegre, v. 35, p. 115-130, vol. Esp., dez. 2016.

⁵⁰⁹ CONSELHO NACIONAL DE JUSTIÇA. Resolução nº 332, de 21 de agosto de 2020.

⁵¹⁰ HABERMAS, Jürgen. **Direito e Democracia**, p. 191.

⁵¹¹ “Como el caso *Rushdie* nos ha recordado, un fundamentalismo que conduzca a una praxis intolerante resulta incompatible con el estado de derecho. Esta praxis se apoya en interpretaciones del mundo — de carácter religioso o propias de una filosofía de la historia — que pretenden exclusividad para una forma de vida privilegiada. Tales concepciones carecen de la conciencia de la falibilidad de sus pretensiones de validez y del respeto por la “carga de la razón” (John Rawls). Por supuesto, las interpretaciones globales del mundo y las convicciones religiosas no tienen que asociarse forzosamente a un falibilismo del tipo que hoy en día acompaña al saber hipotético de las ciencias experimentales. No obstante, las imágenes fundamentalistas del mundo son dogmáticas en el siguiente sentido: no dejan lugar alguno para una reflexión sobre la relación con otras imágenes del mundo con las que comparten el mismo universo de discurso y contra cuyas pretensiones de validez en competencia sólo pueden afirmarse con razones. No dejan lugar alguno para un ‘reasonable disagreement’. Los poderes de la fe subjetivizada del mundo moderno se han caracterizado, por el contrario, por una posición reflexiva que no se limita a aceptar un *modus vivendi* — jurídicamente coactivo bajo las condiciones de la libertad religiosa —. Las imágenes del mundo no fundamentalistas, que Rawls caracteriza como “not unreasonable comprehensive doctrines”, permiten, más bien, en el sentido de la tolerancia de Lessing,

por meio da consideração de parte da cultura de fundo que subsidia as visões de mundo⁵¹² a serem ouvidas com base na alteridade social.⁵¹³

Essa deferência aos aspectos culturais plurais da sociedade demonstra uma influência do pensamento comunitarista (pautado em discriminações qualitativas⁵¹⁴), a qual se volta à acomodação social necessária quando o uso público da racionalidade discursiva se esgota⁵¹⁵, na medida em que as sociedades multiculturais⁵¹⁶ são étnica e religiosamente diversas⁵¹⁷, o que representa um desafio ao uso público da razão.⁵¹⁸

una disputa civilizada entre las diferentes convicciones, en la que una de las partes, sin sacrificar sus propias pretensiones de validez, puede reconocer a las otras partes como contendientes en pro de verdades auténticas" (HABERMAS, Jürgen. *La Lucha por el Reconocimiento en el Estado Democrático de Derecho*, p. 191-192).

⁵¹² *"Indeed, it would be unreasonable to expect churches to teach the virtue of public reasonableness. Public reasonableness is essential in political debate, but is unnecessary and sometimes undesirable in the private sphere. It would be absurd to ask church-goers to abstain from appealing to Scripture III deciding how to run their church"* (KYMLICKA, Will. **Contemporary Political Philosophy: An Introduction**. 2ª ed. Oxford: Oxford University Press, 2002, p. 306).

⁵¹³ "Outra característica da razão pública é que seus limites não se aplicam a nossas deliberações e reflexões pessoais sobre questões políticas, nem à discussão sobre elas por parte de membros de associações como as igrejas e universidades, constituindo tudo isso uma parte vital da cultura de fundo" (Ibid., p. 264).

⁵¹⁴ "Tem havido na filosofia moral moderna uma tendência de sistematização exaustiva. O utilitarismo e o kantismo organizam tudo em torno de uma única razão básica. E, como costuma, acontecer nesses casos, passa a receber crédito entre os proponentes dessas teorias a noção de que a natureza do raciocínio moral é tal que devemos ser capazes de unificar nossas visões morais em torno de uma única base" (TAYLOR, Charles. **As Fontes do Self: A construção da identidade moderna**. Tradução: Adail Ubirajara Sobral e Dinah de Abreu Azevedo. 4ª ed. São Paulo: Edições Loyola, 2013, p. 106-107).

⁵¹⁵ *"At some point, the public reasons may simply run out, and we will be left with conflicting claims based on religious or cultural beliefs that are not publicly shareable. In these circumstances, we need to cultivate the related virtue of accommodation or compromise"* (KYMLICKA, Will. **Contemporary Political Philosophy**, p. 290).

⁵¹⁶ *"En las sociedades multiculturales, la constitución de un estado de derecho sólo puede tolerar aquellas formas de vida que se articulan en el contexto de dichas tradiciones no fundamentalistas, porque la coexistencia en igualdad de derechos de estas formas de vida requiere el reconocimiento recíproco de los diferentes tipos de pertenencia cultural: toda persona debe ser también reconocida como miembro de comunidades integradas cada una en torno a distintas concepciones del bien. La integración ética de grupos y subculturas con sus propias identidades colectivas debe encontrarse, pues, desvinculada del nivel de la integración política, de carácter abstracto, que abarca a todos los ciudadanos en igual medida"* (HABERMAS, Jürgen. *La Lucha por el Reconocimiento en el Estado Democrático de Derecho*, p. 192).

⁵¹⁷ *"This particular conception of public reasonableness-one that requires citizens to consider which of their religious beliefs or cultural traditions are capable of public defence, and to seek honourable compromises when public reasons run out-is distinctly modern. Its prominence in the recent literature on citizenship is partly related to the recognition that modern societies are ethnically and religiously diverse"* (KYMLICKA, Will. **Contemporary Political Philosophy**, p. 290).

⁵¹⁸ *"As I noted last chapter, many liberals are hesitant to adopt this second nation-building level, and prefer to rely solely on people's sense of justice as a motive for accepting the demands of active and responsible citizenship. But in fact, the existence of this sort of nation-building is implicitly assumed by virtually all theorists of deliberative democracy and civic virtue. For example, most accounts of public reasonableness simply take for granted that citizens share a common language, and that democratic states form 'a community of communication'. Indeed, it is difficult to imagine how deliberative democracy is possible without a shared language"* (Ibid., p. 311-312).

Além disso, quando se trata da explicabilidade de decisões judiciais específicas auxiliadas por IA judicial, a argumentação jurídica é o caminho adequado para uma explicitação que faça a mediação entre os motivos exigidos⁵¹⁹ por grupos atomizados⁵²⁰ e as razões compartilhadas na sociedade⁵²¹, pois existem elementos mínimos a caracterizar uma argumentação racional⁵²², afastando-se do mero discurso retórico político ou sectário que não atende às exigências de universalidade⁵²³, de boa-fé argumentativa e de racionalidade.⁵²⁴ Isso é necessário, pois é inevitável, e desejável, que as pessoas tenham visões diferentes.

Embora os cidadãos e o legislativo tenham maior margem de debate público, a explicabilidade de decisões judiciais auxiliadas por IA é parte de uma racionalidade discursiva.⁵²⁵ Ou seja, a explicabilidade traz grande ônus argumentativo que leva o

⁵¹⁹ “As regras da razão permitem a todos problematizar toda afirmação de forma ilimitada. Com isso todo falante pode, como uma criança que sempre continua perguntando mecanicamente “por que?”, colocar todo falante contra a parede. É ainda possível colocar tudo de uma vez como questionável” (ALEXY, Robert. **Teoria Discursiva do Direito**, p. 51).

⁵²⁰ “Entre as razões não-públicas, temos as de todos os tipos de associações: igrejas e universidades, sociedades científicas e grupos profissionais. Como já disse, para agir de forma razoável e responsável, os órgãos coletivos, assim como os indivíduos, precisam de uma forma de argumentação sobre o que deve ser feito. Essa forma de argumentação é pública com respeito a seus membros, mas não-pública com respeito à sociedade política e aos cidadãos em geral. As razões não-públicas compreendem as muitas razões da sociedade civil, em contraste com a cultura política pública. Essas razões são sociais, e certamente não são privadas” (RAWLS, John. **O Liberalismo Político**, p. 269).

⁵²¹ “O princípio da universalização obriga os participantes do discurso a examinar normas controversas, servindo-se de casps particulares *previsivelmente típicos*, para descobrir se elas poderiam encontrar o assentimento refletido de todos os atingidos” (HABERMAS, Jürgen. **Direito e Democracia**, p. 203).

⁵²² “As exigências de igual legitimação, universalidade e ausência de coação podem ser formuladas em três regras. Essas regras correspondem às condições da situação ideal de fala estabelecidas por Habermas.⁵⁹ A primeira regra diz respeito à participação em discursos. Ela tem o seguinte conteúdo:

(2.1) Todo aquele que pode falar pode participar de discursos.

A segunda regra normatiza a liberdade da discussão. Ela pode ser dividida em três exigências:

(2.2) (a) Toda pessoa pode problematizar toda afirmação.

(b) Toda pessoa pode introduzir qualquer afirmação no discurso.

(c) Toda pessoa pode expressar suas opiniões, desejos e necessidades.

(c) possui um significado especial sobretudo em discursos práticos.

A terceira regra tem por fim a tarefa de proteger discursos da força: (2.3) Nenhum falante pode ser impedido, através de coação prevaiente dentro ou fora do discurso, de exercer seus direitos estabelecidos em (2.1) e (2.2)” (ALEXY, Robert. **Teoria Discursiva do Direito**, p. 50).

⁵²³ RAWLS, John. **A Theory of Justice**, p. 17-22.

⁵²⁴ “Isso significa que cada um de nós deve ter e deve estar preparado para explicar um critério acerca de que princípios e diretrizes pensamos que se pode razoavelmente esperar que os outros cidadãos (que também são livres e iguais) subscram junto conosco. Precisamos dispor de algum teste, que estejamos dispostos a explicitar, para dizer quando essa condição é satisfeita” (RAWLS, John. **O Liberalismo Político**, p. 276).

⁵²⁵ “Quem deixa claro aquilo que pretende com a expressão de um juízo de valor ou de dever inclinar-se-á a uma postura cuidadosa e falível. Ele terá claramente em vista sua responsabilidade pessoal pelas ações apoiadas por esses juízos. Também a disposição para a tolerância poderá aumentar. A explicação da pretensão de correção ou de verdade sugerida torna ainda claro que representações concretas normativas bem diferentes podem ter como base o mesmo sistema de regras. As diferenças podem ser explicadas através das diferentes pressuposições fáticas e das diferentes representações

juizador a se guiar pelo uso público da razão.⁵²⁶

A esse respeito, deve-se, mais uma vez, destacar que o esforço da presente obra se pauta na justificação racional de decisões judiciais auxiliadas por Inteligência Artificial, o que se utiliza, neste contexto delimitado, das categorias teóricas de uso público da razão e de consenso delineadas nos pensamentos de Habermas e Alexy. Como se denota do que foi aqui tratado, o objetivo deste tópico se voltou à concepção de explicabilidade da IA judicial e sua relação com a racionalidade necessária ao duplo canal de correção (vide subcapítulo 1.6). Este aviso é necessário, pois não é pretensão deste estudo imprimir uma teoria da Justiça com abrangência geral, motivo pelo qual a exposição de todas as implicações inerentes à concepção de esfera pública habermasiana⁵²⁷ e das críticas às bases filosóficas da teoria não é tarefa a ser exaustivamente aqui tratada.

Após o aviso necessário, destaca-se que a concepção de explicabilidade, como premissa de decisões racionais, demonstra uma distinção importante, eis que ainda que os cidadãos tenham legitimidade para não serem neutros em relação a determinado consenso, os sistemas de Inteligência Artificial judicial podem ser desenvolvidos sob um discurso racional pautado na pretensão de correção⁵²⁸, que se ajusta como base para o desenvolvimento de sistemas de auxílio ao julgamento, pautando-se na função estabilizadora do Judiciário, a qual sempre será desafiada pelas visões pluralistas na fase de supervisão humana dos sistemas de IA judicial voltados ao procedimento de justificação típico do duplo canal de correção. Isto é diferente de se alegar um consenso⁵²⁹ para impedir totalmente o debate democrático sobre determinado tema.

normativas iniciais. Disso se sugere, ao lado da tolerância, a disposição para a discussão" (ALEXY, Robert. **Teoria Discursiva do Direito**, p. 64).

⁵²⁶ HABERMAS, Jürgen. **A Inclusão do Outro**, p. 87.

⁵²⁷ "The public sphere is, then, a locus in which rational views are elaborated which should guide government. This comes to be seen as an essential feature of a free society" (TAYLOR, Charles. **A Secular Age**. Cambridge: Harvard University Press, 2007, p. 189. E-book Kindle).

⁵²⁸ ALEXY, Robert. **Direito, Razão, Discurso**, p. 95-96.

⁵²⁹ "The outcome of this pluralism and mutual fragilization will often be a retreat of religion from the public square. In one way this is inevitable and in the circumstances good. Justice requires that a modern democracy keep an equal distance from different faith positions. The language of some public bodies, for instance courts, has to be free from premises drawn from one or other position. Our cohesion depends on a political ethic, democracy and human rights, basically drawn from the Modern Moral Order, to which different faith and non - faith communities subscribe, each for their own divergent reason. We live in a world of what John Rawls has described as 'overlapping consensus'. But in another way, as José Casanova has argued, religious discourse will be very much in the public square. Democracy requires that each citizen or group of citizens speak the language in public debate that is most meaningful to them. Prudence may urge us to put things in terms which others relate to, but to

Logo, partindo das premissas de Habermas e Alexy, fica caracterizado que a racionalidade argumentativa é o caminho para o desenvolvimento de sistemas de Inteligência Artificial judicial baseados no uso público da razão, que nasce da percepção que não existe uma versão unificada de mundo.⁵³⁰ Em suma, as teorias de Habermas e Alexy têm pontos de ligação inerentes à racionalidade argumentativa que não se fecha para o pluralismo.

Portanto, os sistemas de IA para auxílio ao julgamento devem ser pensados a partir da premissa de que o consenso do processo (numa dialética limitada e controlada a cada caso topicamente considerado) não supera a racionalidade discursiva formada com o auxílio do duplo canal de correção inerente à interação comunicativa humano-máquina. Dessa maneira, o modelo de supervisão humana pelo julgador do caso (art. 19, Parágrafo único, da Resolução 332/2020 do CNJ⁵³¹) trará o equilíbrio necessário ao desenvolvimento e ao funcionamento de sistemas de Inteligência Artificial judicial.

Foi observado que a explicabilidade das decisões não é um problema totalmente novo, mas se trata de questão tradicional acerca da transparência e da sinceridade na construção das decisões judiciais.

Portanto, o direito à explicabilidade das decisões judiciais tomadas com auxílio de Inteligência Artificial é fomentado através da definição de um espaço democrático de argumentação marcada pela pluralidade.

Evidencia-se que a Inteligência Artificial judicial, ainda que esteja experimentando recente abordagem regulatória (Conselho Nacional de Justiça, Lei Geral de Proteção de Dados, União Europeia, OCDE, UNESCO, etc. – vide os

require this would be an intolerable imposition on citizen speech" (TAYLOR, Charles. **A Secular Age**, p. 532).

⁵³⁰ "Seria um equívoco concluir, em virtude do fato de as regras do discurso não garantirem a certeza em todos os casos, que qualquer coisa é compatível com elas. Há alguns resultados que, através delas, são forçosamente exigidos ou excluídos. Isso vale por exemplo para normas que excluem inteiramente pessoas do desenvolvimento discursivo, atribuindo a elas o status jurídico de escravas. Nesse sentido pode-se falar em 'impossibilidade discursiva' e em 'necessidade discursiva'.

Nos casos em que dois enunciados normativos incompatíveis entre si podem ser fundamentados sem violar as regras do discurso pode-se falar em 'possibilidade discursiva'. O espaço do discursivamente possível é na verdade amplo, mas não chega nem perto de significar ser tudo possível. Ele é ainda limitado pelo fato de que muitas coisas que seriam possíveis de acordo com as regras do discurso sozinhas, ou seja, sob os pressupostos de indivíduos em diversas condições, não são possíveis de acordo com as regras do discurso em conjunto com as condições fáticas dos participantes do discurso. Como, nesse sentido, a teoria do discurso deixa muito àqueles que discutem, ela pressupõe que eles são suficientemente razoáveis para, sob as condições do discurso, alcançar resultados razoáveis" (ALEXY, Robert. **Teoria Discursiva do Direito**, p. 62).

⁵³¹ CONSELHO NACIONAL DE JUSTIÇA. Resolução nº 332, de 21 de agosto de 2020.

subcapítulos 2.7 e 2.8), demanda um estudo típico da filosofia e da teoria do Direito, a fim de serem sedimentadas as bases para o reconhecimento do direito fundamental à transparência das decisões tomadas com base em Inteligência Artificial.

Nesse caminho, evidenciou-se que a nova camada de linguagem da decisão judicial, implementada através da Inteligência Artificial, traz consigo riscos (opacidade e vieses algorítmicos, por exemplo), que devem ser tratados na computação aplicada ao direito mediante soluções híbridas de mútua cooperação científica entre teoria da Justiça e Ciência da Computação, o que é possível mediante arquiteturas de sistemas voltados para a explicabilidade de decisões tomadas por IA.

Nesse aspecto, o uso público da razão aplicado à decisão judicial tem elementos que podem ser utilizados para potencializar as técnicas já existentes de Inteligência Artificial explicável (xAI), desde o processo de desenvolvimento dos algoritmos e modelos de IA até os resultados de processamento computacional.

Dessa forma, uma IA judicial pode empregar as abordagens procedimental, de representação do conhecimento ou de *machine learning* (incluindo-se o dilúvio de dados necessário aos modelos *deep learning*) em relação à sua programação. Todavia, deve-se sempre desenvolver os sistemas com a possibilidade de explicabilidade compreensível aos humanos.

Nesse ponto, até mesmo para que adequadamente existam desenvolvimento, controle e supervisão de sistemas de IA judicial, é um caminho natural que os profissionais do Direito do futuro (e do presente) passem da figura do operador do direito para posições mais próximas do que se pode chamar de desenvolvedor de software jurídico.

A crítica ao entrelaçamento entre direito e Inteligência Artificial continuará a ter sua relevância, mas, se ficar pousada em uma posição exclusivamente reacionária, em pouco tempo será vencida pela realidade tecnológica que se impõe.

Portanto, a IA de auxílio à decisão judicial (pensamento/processamento híbrido humano-máquina) já funciona com as limitações que a tecnologia impõe, cabendo desde já reconhecer a explicabilidade da IA jurídica como um direito necessário para uma sociedade plural e justa.

2.5 *Due xAI of Law*: o Direito Fundamental à Explicabilidade

É provável que o título chame a atenção do leitor. Todavia, *Due xAI of Law* naturalmente faz referência ao *Due Process of Law* (devido processo legal), com a aposição do acrônimo *xAI* (*explainable artificial intelligence*), em lugar da palavra *process*. O título é uma provocação teórica inerente à inclusão de Inteligência Artificial explicável no âmbito de processos judiciais, conforme o trecho a seguir passa a tratar.

A democracia e a sociedade evoluíram no sentido de que as decisões de caráter público devem ser tomadas de maneira a possibilitar o controle pelos interessados, o que implica num caráter formal referente à publicação dos atos estatais e num plano material da existência de uma fundamentação lógico-racional para a decisão⁵³².

O direito atua criando os estímulos à adequada e racional fundamentação da decisão judicial⁵³³, bem como demarca limitações aos eventuais desbordes da racionalidade causados por parte do julgador.

Nessa linha, no aspecto normativo a transparência passou de uma mera formalidade, pautada numa exigência de publicidade formal, para o centro da arquitetura de decisões⁵³⁴ voltada à política pública de transparência estatal, cujas ações devem ser marcadas pela proatividade na explicação dos motivos pelos quais as decisões são tomadas, comportamento ativo que caracteriza um padrão⁵³⁵ que é estimulado pelo sistema jurídico.

Dessa forma, existe um coeso aglomerado normativo⁵³⁶ que indica que a publicidade de atos estatais, incluindo nesse conceito as decisões judiciais, é um valor⁵³⁷ que envolve o sistema jurídico-constitucional do Brasil, haja vista as previsões do art. 5º, XXXIII e XXXIV, “b”, do art. 37, *caput*, e do art. 93, IX, da Constituição Federal, do art. 8º, do art. 11 e do art. 489, §1º, do Código de Processo Civil, da Lei 9.051/1995, da Lei 12.527/2011 (Lei de Acesso à Informação), do art. 50 da Lei

⁵³² ATIENZA, Manuel. **As Razões do Direito**: Teoria da Argumentação Jurídica. Tradução: Maria Cristina Guimarães Cupertino. 2ª ed. Rio de Janeiro: Forense Universitária, 2016, p. 268.

⁵³³ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 170-171.

⁵³⁴ THALER, Richard H.; SUNSTEIN, Cass R. Op. cit., posição 89-90.

⁵³⁵ “*Drawing on the underlying research, both governments and the private sector are becoming far more sophisticated in using default rules to promote their preferred outcomes — sometimes for excellent reasons, sometimes out of self-interest*” (SUNSTEIN, Cass R. **Choosing Not to Choose**. Oxford: Oxford University Press, 2015, p. 71. E-book Kindle).

⁵³⁶ ÁVILA, Humberto Bergmann. Op. cit., p. 49.

⁵³⁷ ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 147.

9.784/1999 e do art. 4º da Lei 11.419/2006 (Lei de Processo Judicial Eletrônico). Assim, há um bloco normativo coeso que demonstra que o chamado princípio da publicidade dos atos estatais evoluiu para o dever de transparência dos órgãos estatais. Logo, a construção de um ambiente institucional favorável à *accountability* decorre do “fomento ao desenvolvimento da cultura de transparência” (art. 3º, IV, da Lei 12.527/2011). O referido bloco normativo se alinha “a essa tendência de publicidade ampla a reger as atividades da Administração, invertendo a regra do segredo e do oculto que predominava”⁵³⁸.

Nesse sentido, em regulamentação ao direito fundamental à proteção de dados pessoais (art. 5º, LXXIX, da Constituição Federal), o art. 6º da Lei 13.709/2018 (Lei Geral de Proteção de Dados Pessoais) enumera os princípios inerentes à proteção de dados no Brasil, e o inciso VI traz o conceito do princípio de **transparência** como a “garantia, aos titulares, de informações claras, precisas e facilmente acessíveis sobre a realização do tratamento e os respectivos agentes de tratamento, observados os segredos comercial e industrial”. Esse princípio encontra eco nos princípios da finalidade (inciso I); do livre acesso (inciso IV); da qualidade dos dados (inciso V); da não discriminação (inciso IX); e da responsabilização e prestação de contas (inciso X); bem como no dever de transparência do controlador (art. 10, §2º). Cabe um destaque ao art. 20, §1º, da Lei 13.709/2018 ao prever que o “controlador deverá fornecer, sempre que solicitadas, informações claras e adequadas a respeito dos critérios e dos procedimentos utilizados para a decisão automatizada, observados os segredos comercial e industrial” (destaque não original).

Desse modo, num dos passos mais atuais sobre o tema no Brasil, o art. 20 da Lei Geral de Proteção de Dados Pessoais (Lei 13.709/2018) prevê: a) o direito à revisão humana das decisões tomadas de maneira automatizada que envolvam dados pessoais (*caput*); b) o direito ao recebimento de informações claras e adequadas acerca dos critérios e dos procedimentos utilizados para a decisão automatizada, desde que respeitados os segredos comercial e industrial (§1º); e c) o direito à solicitação de auditoria da autoridade nacional de proteção de dados em relação aos eventuais aspectos discriminatórios advindos de decisões automatizadas (Inteligência

⁵³⁸ MEDAUAR, Odete. **Direito Administrativo Moderno**. 19ª ed. São Paulo: Revista dos Tribunais, 2015, p. 159.

Artificial) tomadas por sistemas aos quais se aplique o segredo comercial ou industrial (§2º).

Nessa linha, uma visão de otimização da aplicabilidade do art. 20 da Lei 13.709/2018 às decisões judiciais tomadas com auxílio de sistemas de Inteligência Artificial⁵³⁹ pode ser implementada como forma de garantir a efetiva transparência do exercício do Poder Judiciário, fomentada pelo art. 93, IX, da Constituição Federal, o qual determina que “todos os julgamentos dos órgãos do Poder Judiciário serão públicos, e fundamentadas todas as decisões, sob pena de nulidade...” (destaque não original) e pelo direito fundamental à proteção de dados pessoais (art. 5º, LXXIX, da Constituição Federal).

Medidas que fomentam a publicidade das decisões judiciais e o direito fundamental à proteção de dados foram adotadas pela Resolução 332, de 21 de agosto de 2020⁵⁴⁰, do Conselho Nacional de Justiça (CNJ), a qual prevê, em seu art. 8º, que a transparência da IA judicial consiste: a) na divulgação responsável (inciso I); b) na indicação dos objetivos e resultados pretendidos; c) na documentação dos riscos identificados e indicação dos instrumentos de segurança da informação e controle (inciso III); d) na identificação do motivo em caso de dano (inciso IV); e) na existência de mecanismos de auditoria e certificação de boas práticas (inciso V); e f) no **“fornecimento de explicação satisfatória e passível de auditoria por autoridade humana quanto a qualquer proposta de decisão apresentada pelo modelo de Inteligência Artificial, especialmente quando essa for de natureza judicial”**⁵⁴¹ (inciso VI – destaque nosso). Destaca-se ainda que o art. 19, *caput*, da Resolução 332/2020 do CNJ determina que os “sistemas computacionais que utilizem modelos de Inteligência Artificial como ferramenta auxiliar para a elaboração de decisão judicial **observarão, como critério preponderante para definir a técnica utilizada, a explicação dos passos que conduziram ao resultado**”⁵⁴² (destaque nosso). Por fim, o art. 25, *caput*, da referida Resolução prevê que qualquer IA judicial “deverá assegurar **total transparência na prestação de contas**”⁵⁴³ (destaque nosso).

⁵³⁹ Aqui se utiliza o conceito amplo de inteligência artificial, o que passa desde sistemas de busca simples e sistemas baseados no conhecimento (sistemas especialistas) até sistemas com *machine learning* e *deep learning*.

⁵⁴⁰ CONSELHO NACIONAL DE JUSTIÇA. Resolução nº 332, de 21 de agosto de 2020.

⁵⁴¹ Id.

⁵⁴² Id.

⁵⁴³ Id.

Dessa forma, como pressuposto teórico para se falar de modelos de Inteligência Artificial explicável (xAI)⁵⁴⁴ é necessário que sejam traçadas as bases do que se pode entender como o sistema de normas (regras e princípios) aplicáveis à busca de *accountability*⁵⁴⁵ *by design*⁵⁴⁶ por meio de maiores graus de transparência, interpretabilidade e explicabilidade dos algoritmos de IA, inclusive através dos conceitos trazidos pelas recomendações e declarações de organismos internacionais (ainda que sem caráter obrigatório no plano das relações jurídicas internas), de maneira a delinear as políticas públicas sobre o tema.

Sendo assim, a melhora na transparência⁵⁴⁷ é a condição ético-normativa para o desenvolvimento da IA na sociedade, especialmente dos sistemas de Inteligência Artificial aplicados às decisões judiciais. Destaca-se que na obra *Eu, Robô* de Isaac Azimov foram descritas as três leis da robótica⁵⁴⁸, evidenciando que a necessidade de explicabilidade pode ser inferida da aplicação da Segunda Lei: “um robô deve obedecer às ordens dadas por seres humanos”⁵⁴⁹, exceto para cumprir a Primeira Lei⁵⁵⁰. Inclusive, em tal sentido, Frank Pasquale⁵⁵¹ defende a explicabilidade (*explainability*) como a quarta lei da robótica. Em suma, sem explicabilidade não há como se provar que a “ordem” humana foi obedecida pela Inteligência Artificial.

Nesse caminho, o escrutínio público das razões de uma decisão⁵⁵² é parte dos elementos que integram a tomada de decisões racionais⁵⁵³ no plano democrático. A

⁵⁴⁴ “xAI encompasses a range of efforts to explain — or help humans interpret — how a particular machine learning model reached its conclusion. The concept of an explanation here ‘has come to refer to providing insight into the internal state of an algorithm, or to human-understandable approximations of the algorithm.’ xAI provides a variety of benefits: It can foster trust between humans and the system, identify cases in which the system appears to be biased or unfair, and bolster our own knowledge of how the world works. As discussed below, in legal settings xAI can benefit judges who wish to rely on the algorithms for decisional support, litigants who seek to persuade judges that their use of algorithms is defensible, and defendants who wish to challenge predictions about their dangerousness” (DEEKS, Ashley. Op. cit.).

⁵⁴⁵ CAMPOS, Anna Maria. Accountability: quando poderemos traduzi-la para o português? **Revista de Administração Pública**, Rio de Janeiro, v. 24, n. 2, p. 30-50, jun. 1990.

⁵⁴⁶ WAGNER, Ben. Op. cit.

⁵⁴⁷ BOSTROM, Nick. Op. cit., posição 5334-5342.

⁵⁴⁸ ASIMOV, Isaac. Op. cit., posição 686-690.

⁵⁴⁹ Ibid., posição 687-688.

⁵⁵⁰ A Primeira Lei da robótica prevê que: “um robô não pode ferir um ser humano ou, por inação, permitir que um ser humano venha a ser ferido” (Ibid., posição 686-687).

⁵⁵¹ “One key element of explainability is a clear sense of the history of a robot — how was it first programmed, to what has it been exposed, and how has this interplay between hardware, software, and the external environment resulted in present behavior” (PASQUALE, Frank. Toward a Fourth Law of Robotics: Preserving Attribution, Responsibility, and Explainability in an Algorithmic Society. **Ohio State Law Journal**, Vol. 78, 2017).

⁵⁵² HABERMAS, Jürgen. **A Inclusão do Outro**, p. 79.

⁵⁵³ “A primeira regra refere-se à admissão no discurso. O conteúdo é o seguinte:

transparência dos atos estatais é um elemento básico para a construção de um ambiente institucional favorável à *accountability* algorítmica⁵⁵⁴ e ao debate público como influenciadores de melhores razões para as decisões.

A respeito do tema, em 2017 foi realizada a *Asilomar Conference on Beneficial AI*, evento científico no qual os membros do *Future of Life Institute* definiram os vinte e três *Asilomar AI Principles*⁵⁵⁵ como padrões éticos a serem seguidos no desenvolvimento da Inteligência Artificial, destacando no presente tópico a relação entre a IA jurídica e os princípios da segurança (*safety*), da transparência de falhas (*failure transparency*), da transparência judicial (*judicial transparency*), da responsabilidade (*responsibility*), do alinhamento de valores (*value alignment*), do controle humano (*human control*) e da não subversão da IA (*non-subversion*).

Os Princípios de Asilomar representam um engajamento inicial da comunidade científica no debate acerca dos padrões éticos e dos limites jurídicos para implementação da Inteligência Artificial na sociedade, com destaque para a indicação do necessário controle humano sobre as decisões tomadas por meio de IA⁵⁵⁶ e para a adequada transparência de sistemas de IA utilizados pelo Poder Judiciário (princípio da transparência judicial), especialmente no que se refere à possibilidade do controle de falhas cometidas pela AI jurídica (princípio da transparência de falhas)⁵⁵⁷.

Nesse caminho, o **princípio da transparência judicial** define que “qualquer envolvimento de um sistema autônomo na tomada de decisões judiciais deve fornecer

(2.1.) Quem pode falar, pode tomar parte no discurso.

A segunda regra regula a liberdade de discussão. Pode subdividir-se em três exigências:

(2.2) (a) Todos podem problematizar qualquer asserção.

(b) Todos podem introduzir qualquer asserção no discurso.

(c) Todos podem expressar suas opiniões, desejos e necessidades”.

(...)

“Porém, essas regras não são sem sentido. Os argumentos que não foram aceitos se cumpridos (2.1) a (2.3) (incluindo-se as demais regras do discurso) devem considerar-se não válidos. Desta maneira, formam, junto às demais regras do discurso, um *critério hipotético-normativo* para a correção das proposições normativas” (ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 172).

⁵⁵⁴ PASQUALE, Frank. Data-Informed Duties in AI Development. **Columbia Law Review**, v. 119, n. 7, p. 1917-1940, 2019.

⁵⁵⁵ FUTURE OF LIFE INSTITUTE. Op. cit.

⁵⁵⁶ Nesse sentido, o art. 19, Parágrafo único, da Resolução 332/2020 do CNJ prevê: “Os sistemas computacionais com atuação indicada no caput deste artigo deverão permitir a supervisão do magistrado competente” (CONSELHO NACIONAL DE JUSTIÇA. Resolução nº 332, de 21 de agosto de 2020).

⁵⁵⁷ O art. 25, Parágrafo único, V, da Res. 332/2020 do CNJ indica que o dever de prestação de contas inclui “a demonstração de efetiva publicidade quanto à natureza do serviço oferecido, técnicas utilizadas, desempenho do sistema e riscos de erros” (Id.).

uma explicação satisfatória auditável por uma autoridade humana competente”⁵⁵⁸, sendo que o **princípio da transparência de falhas** prevê que “se um sistema de IA causar danos, deve ser possível verificar o porquê”⁵⁵⁹. Sendo assim, percebe-se nos Princípios de Asilomar uma sensível preocupação com a tomada de decisão judicial por meio de Inteligência Artificial, o que representa uma demanda melhora nos níveis de transparência, supervisão e controle⁵⁶⁰ da IA. Todavia, não se coloca um impeditivo técnico para a utilização de Inteligência Artificial pelo Poder Judiciário.

Nesse sentido, o crescimento da utilização da Inteligência Artificial na sociedade gerou reações junto às organizações internacionais, como os *Ethical Guidelines for Trustworthy AI*⁵⁶¹ da União Europeia, os *G20 AI Principles*⁵⁶², o *G7’s Charlevoix Common Vision for the Future of AI*⁵⁶³, os *AI Principles* do Departamento de Defesa dos EUA⁵⁶⁴, a *OECD’s Recommendation of the Council on AI*⁵⁶⁵ e a *Recommendation on the Ethics of Artificial Intelligence* da UNESCO⁵⁶⁶.

As iniciativas para buscar padrões éticos para o desenvolvimento da Inteligência Artificial são diversas, e muito do que é traçado tem linhas comuns no respeito à humanidade e nos limites de segurança no uso da IA. O presente tópico faz o recorte da pesquisa especificamente na Recomendação 449 da OCDE e na Recomendação sobre IA da UNESCO.

Na Recomendação 449 da OCDE (Organização para a Cooperação e Desenvolvimento Econômico), cuja adesão foi manifestada pelo Brasil em 21/05/2019, destaca-se a seção 1.3, que trata da transparência e da explicabilidade (*transparency and explainability*)⁵⁶⁷ dos sistemas de Inteligência Artificial.

⁵⁵⁸ Tradução livre de: “Any involvement by an autonomous system in judicial decision-making should provide a satisfactory explanation auditable by a competent human authority” (FUTURE OF LIFE INSTITUTE. Op. cit.).

⁵⁵⁹ Tradução livre de: “If an AI system causes harm, it should be possible to ascertain why” (Id.).

⁵⁶⁰ RUSSELL, Stuart; DEWEY, Daniel; TEGMARK, Max. Op. cit.

⁵⁶¹ UNIÃO EUROPEIA. **Ethics Guidelines for Trustworthy AI**.

⁵⁶² G20. **G20 AI Principles**. [2019]. Disponível em: <https://www.g20-insights.org/wp-content/uploads/2019/07/G20-Japan-AI-Principles.pdf>. Acesso em: 25 jul. 2020.

⁵⁶³ G7. **G7’s Charlevoix Common Vision for the Future of AI**. [2018]. Disponível em: <https://www.mofa.go.jp/mofaj/files/000373837.pdf>. Acesso em: 25 jul. 2020.

⁵⁶⁴ ESTADOS UNIDOS DA AMÉRICA. **AI Principles: Recommendations on the Ethical Use of Artificial Intelligence by the Department of Defense**. [2019]. Disponível em: https://media.defense.gov/2019/Oct/31/2002204458/-1/-1/0/DIB_AI_PRINCIPLES_PRIMARY_DOCUMENT.PDF. Acesso em: 08 ago. 2020.

⁵⁶⁵ OCDE. Op. cit.

⁵⁶⁶ UNESCO. Op. cit.

⁵⁶⁷ OCDE. Op. cit.

Nesse aspecto, na referida seção 1.3 a redação foi definida de maneira que os atores em Inteligência Artificial devem se comprometer com a transparência e a divulgação responsável dos sistemas de IA, no sentido de: a) **promover um entendimento geral dos sistemas de IA** (inciso I); b) **conscientizar as partes interessadas sobre suas interações com os sistemas de IA**, inclusive no local de trabalho (inciso II); c) **habilitar as pessoas afetadas a entender o resultado** da solução dada através da IA (inciso III); e d) possibilitar o **desafio dos resultados de um sistema de IA pelas pessoas afetadas de maneira adversa**, mediante **informações claras e fáceis de entender** sobre os fatores e sobre a lógica que serviu de base para a predição, a recomendação ou a decisão (inciso IV).

Além disso, a seção 1.5⁵⁶⁸ da Recomendação 449 da OCDE prevê um dever de *accountability*, pois os atores em Inteligência Artificial devem ser responsáveis pelo adequado funcionamento dos sistemas de IA e pelo respeito aos princípios de transparência e explicabilidade, com base em seus papéis, no contexto da IA na sociedade e em consistência com o estado da arte da Ciência da Computação.

Assim, a Recomendação 449 da OCDE traz avanços no sentido da futura construção de sistema normativo nacional de regulação da IA⁵⁶⁹, com especial enfoque para os direitos à transparência e à explicabilidade dos algoritmos de IA. Vê-se, todavia, que a seção 1.5 da Recomendação 449 reconhece os limites fáticos da efetivação dos direitos de transparência e explicabilidade, uma vez que remete ao “estado da arte” da Inteligência Artificial como o padrão a ser seguido, o que garante

⁵⁶⁸ “1.5.Accountability

AI actors should be accountable for the proper functioning of AI systems and for the respect of the above principles, based on their roles, the context, and consistent with the state of art”. (Id.)

⁵⁶⁹ O que passa pelo Projeto de Lei 21/2020 da Câmara dos Deputados, que atualmente tramita no Senado Federal. O art. 6º, V, do Projeto de Lei 21/2020 tem a proposta de definição do princípio de transparência:

“V – transparência: direito das pessoas de serem informadas de maneira clara, acessível e precisa sobre a utilização das soluções de inteligência artificial, salvo disposição legal em sentido contrário e observados os segredos comercial e industrial, nas seguintes hipóteses:

a) sobre o fato de estarem se comunicando diretamente com sistemas de inteligência artificial, tal como por meio de robôs de conversação para atendimento personalizado on-line (*chatbot*), quando estiverem utilizando esses sistemas;

b) sobre a identidade da pessoa natural, quando ela operar o sistema de maneira autônoma e individual, ou da pessoa jurídica responsável pela operação dos sistemas de inteligência artificial;

c) sobre critérios gerais que orientam o funcionamento do sistema de inteligência artificial, assegurados os segredos comercial e industrial, quando houver potencial de risco relevante para os direitos fundamentais” (BRASIL. Senado Federal. **Projeto de Lei nº 21, de 2020**. Estabelece fundamentos, princípios e diretrizes para o desenvolvimento e a aplicação da inteligência artificial no Brasil; e dá outras providências. Brasília. Senado Federal, 03 fev. 2022. Disponível em: legis.senado.leg.br/sdleg-getter/documento?dm=9063365&ts=1656528542410&disposition=inline. Acesso em: 17 jul. 2022).

o direito ao melhor grau de transparência/explicabilidade possível pelas técnicas atualmente existentes. Ao mesmo tempo, a seção 1.5 reconhece implicitamente que o estado da técnica em IA tem limites insuperáveis na atualidade, o que levará a existência de previsões, recomendações e decisões realizadas por IA que, no atual “estado da arte”, continuarão a possuir baixo nível de transparência (decisões tomadas por redes neurais⁵⁷⁰, por exemplo).

Vê-se que a Recomendação 449 da OCDE faz um enquadramento inicial dos desafios de transparência e explicabilidade dos sistemas de IA.

2.6 Inteligência Artificial e Transparência *by Design*: a Explicabilidade Necessária e a Interpretabilidade Possível

A distinção entre explicabilidade e interpretabilidade é necessária, a fim de se aprofundar nos efeitos jurídicos dessa caracterização de ambos os institutos.

Nessa linha, a UNESCO elaborou a *Recommendation on the Ethics of Artificial Intelligence*, documento que tem como escopo o estabelecimento de padrões para a ética e para as políticas públicas no desenvolvimento da Inteligência Artificial. Na Recomendação busca-se a formulação de padrões éticos que alcancem todo o ciclo de vida dos sistemas de IA no que se refere à pesquisa, ao projeto, ao desenvolvimento, à implantação e à utilização (inclusive no se refere à manutenção, à operação, ao final de ciclo de uso e ao desligamento).⁵⁷¹

Ademais, pode ser destacada seção 3, “b”, da referida Recomendação, que prevê que mesmo nas ciências sociais e humanas a Inteligência Artificial tem “implicações para nossos conceitos de entendimento e explicação científicos, e cria uma nova base para a tomada de decisões”.⁵⁷²

Dessa maneira, os modelos de IA desenvolvidos para auxílio na tomada de decisões judiciais são afetados na Recomendação da UNESCO pelos seguintes

⁵⁷⁰ “One common viewpoint in the deep neural network community is that the level of interpretability and theoretical understanding needed to for transparent explanations of large DNNs remains out of reach” (GILPIN, Leilani H. et al. Op. cit.).

⁵⁷¹ UNESCO. Op. cit., p. 6-8.

⁵⁷² Tradução livre de: “(b) Science, in the broadest sense and including all academic fields from the natural sciences and medical sciences to the social sciences and humanities, as AI technologies bring new research capacities and approaches, have implications for our concepts of scientific understanding and explanation, and create a new basis for decision-making” (Ibid., p. 4).

princípios: proporcionalidade (*proportionality*); justiça e não discriminação (*fairness and non-discrimination*)⁵⁷³; supervisão e determinação humana (*human oversight and determination*); transparência (*transparency*); explicabilidade (*explainability*); responsabilidade (*responsibility*); e *accountability*. Tais princípios devem ser interpretados de maneira a garantir, principalmente, as ações voltadas às áreas de intervenção (*policies areas*) inerentes: à avaliação do impacto ético (*Policy Area 1*⁵⁷⁴), à governança e à gestão éticas (*Policy Area 2*) e à comunicação e à informação (*Policy Area 9*).

Em relação à **proporcionalidade** (*proportionality*) a recomendação da UNESCO estabelece que os processos relacionados ao ciclo de vida dos sistemas de Inteligência Artificial não devem ultrapassar o necessário para se alcançar as metas ou os objetivos que sejam legítimos, bem como cada sistema de IA deve ser apropriado ao contexto em que está inserido.⁵⁷⁵ Desse modo, deve existir uma prévia justificativa a respeito do uso de sistemas de IA e dos métodos de Inteligência Artificial utilizados, mediante a demonstração de: **a)** que o método de IA escolhido é apropriado e proporcional para alcançar um determinado objetivo legítimo; **b)** que o método de Inteligência Artificial utilizado não viola os valores previstos na Recomendação da UNESCO (especialmente a proteção aos direitos humanos); **c)** que existe adequação do método de IA utilizado ao contexto; e **d)** que a Inteligência Artificial se baseia em fundamentos científicos rigorosos.⁵⁷⁶ Com base na proporcionalidade, a seção 26 da recomendação da UNESCO ainda estabelece que deve ser aplicada “determinação

⁵⁷³ “28. AI actors should promote social justice and safeguard fairness and non-discrimination of any kind in compliance with international law. This implies an inclusive approach to ensuring that the benefits of AI technologies are available and accessible to all, taking into consideration the specific needs of different age groups, cultural systems, different language groups, persons with disabilities, girls and women, and disadvantaged, marginalized and vulnerable people or people in vulnerable situations. Member States should work to promote inclusive access for all, including local communities, to AI systems with locally relevant content and services, and with respect for multilingualism and cultural diversity” (Ibid., p. 7-8).

⁵⁷⁴ “50. Member States should introduce frameworks for impact assessments, such as ethical impact assessment, to identify and assess benefits, concerns and risks of AI systems, as well as appropriate risk prevention, mitigation and monitoring measures, among other assurance mechanisms. Such impact assessments should identify impacts on human rights and fundamental freedoms, in particular but not limited to the rights of marginalized and vulnerable people or people in vulnerable situations, labour rights, the environment and ecosystems and ethical and social implications, and facilitate citizen participation in line with the values and principles set forth in this Recommendation” (Ibid., p. 10-11).

⁵⁷⁵ Ibid., p. 7.

⁵⁷⁶ Id.

humana final”⁵⁷⁷ nos sistemas de IA em que as decisões possam ter um impacto irreversível, de difícil reversibilidade ou em situações de “vida ou morte”.⁵⁷⁸

Acerca da **justiça e não discriminação** (*fairness and non-discrimination*) o documento elaborado pela UNESCO prevê que as decisões tomadas por meio de Inteligência Artificial devem buscar padrões de justiça com base na equidade, na igualdade e na inclusão, a fim de se esforçar para minimizar e evitar o reforço ou a perpetuação de vieses, incluindo os raciais, étnicos, de gênero, de idade e culturais.⁵⁷⁹

A **supervisão e determinação humana** (*human oversight and determination*) consiste na obrigação, relacionada a qualquer fase do ciclo de vida dos sistemas de Inteligência Artificial, de que seja sempre possível atribuir a responsabilidade, nos âmbitos ético e de responsabilidade legal (em especial a responsabilidade civil), a pessoas físicas e jurídicas existentes.⁵⁸⁰ A partir disso, se qualifica uma abordagem integrativa entre a supervisão humana individual e a supervisão pública inclusiva.⁵⁸¹ Desse modo, por razões de eficácia os humanos (e a sociedade) podem escolher confiar em sistemas baseados em Inteligência Artificial, pois apenas aos humanos cabe essa opção em contextos limitados, usando a IA como um auxiliar no processo humano de decisão, uma vez que uma IA não pode substituir a responsabilidade e a *accountability* humanas.⁵⁸² A partir do princípio da supervisão e determinação humanas também se reitera que as “decisões de vida e morte não devem ser cedidas a sistemas de IA”.⁵⁸³ Assim, segundo a seção 63 da recomendação da UNESCO, nos sistemas de IA utilizados pelo Poder Judiciário devem existir salvaguardas suficientes para que o princípio da supervisão humana seja garantido.⁵⁸⁴

A **transparência** (*transparency*) permite que as pessoas entendam como os sistemas de Inteligência Artificial são pesquisados, projetados, desenvolvidos, implantados e utilizados (conforme cada contexto e levando-se em conta a

⁵⁷⁷ Id.

⁵⁷⁸ “In scenarios where decisions are understood to have an impact that is irreversible or difficult to reverse or may involve life and death decisions, final human determination should apply. In particular, AI systems should not be used for social scoring or mass surveillance purposes” (Id).

⁵⁷⁹ Ibid., p. 7-8.

⁵⁸⁰ Ibid., p. 8.

⁵⁸¹ Id.

⁵⁸² Ibid., p. 8-9.

⁵⁸³ Tradução livre de: “As a rule, life and death decisions should not be ceded to AI systems” (Ibid., p. 9).

⁵⁸⁴ “In case AI systems are used by the judiciary, sufficient safeguards are needed to guarantee inter alia the protection of fundamental human rights, the rule of law, judicial independence as well as the principle of human oversight, and to ensure a trustworthy, public interest-oriented and human-centric development and use of AI systems in the judiciary” (Ibid., p. 12).

sensibilidade de cada sistema de IA), o que inclui informações adequadas sobre os fatores que afetam uma previsão ou uma decisão específica.⁵⁸⁵ Na abordagem da UNESCO, a transparência não precisa, a princípio, incluir o compartilhamento de códigos ou conjuntos de dados específicos⁵⁸⁶, uma vez que a transparência é uma questão sócio-técnica, a fim de se melhorar a confiança dos seres humanos em relação aos sistemas de Inteligência Artificial.⁵⁸⁷ Contudo, em casos de grave ameaça aos direitos humanos, a transparência pode envolver o acesso ao código-fonte ou à base de dados utilizada no treinamento da IA.⁵⁸⁸ A esse respeito o art. 18, *caput*, da Resolução 332, de 21 de agosto de 2020, do Conselho Nacional de Justiça (CNJ) prevê que “Os usuários externos devem ser informados, em linguagem clara e precisa, quanto à utilização de sistema inteligente nos serviços que lhes forem prestados”.⁵⁸⁹

Já a **explicabilidade** (*explainability*)⁵⁹⁰ se pauta no dever de buscar tornar inteligíveis aos seres humanos os resultados dos sistemas de Inteligência Artificial, através do fornecimento de informações sobre esses, o que leva a uma melhor compreensão das entradas (*inputs*), das saídas (*outputs*) e do comportamento de cada componente fundamental algorítmico que contribui para o resultado do processamento.⁵⁹¹

Portanto, explicabilidade e transparência estão intimamente ligadas, haja vista que os resultados e o processamento que levam a essas devem ser compreensíveis e rastreáveis dentro de seu contexto de utilização. Nessa linha, o art. 8º, VI, da Resolução 332/2020 do CNJ, ao tratar dos elementos da transparência para a IA judicial, prevê o direito do cidadão ao acesso à “explicação satisfatória e passível de

⁵⁸⁵ Ibid., p. 9.

⁵⁸⁶ Contudo, o art. 10, III, da Res. 332/2020 do CNJ determina que os Tribunais depositem naquele órgão de controle cada modelo de Inteligência Artificial que vier a ser desenvolvido (CONSELHO NACIONAL DE JUSTIÇA. Resolução nº 332, de 21 de agosto de 2020).

⁵⁸⁷ UNESCO. Op. cit., p. 9.

⁵⁸⁸ “In cases of serious threats of adverse human rights impacts, transparency may also require the sharing of code or datasets” (Id.)

⁵⁸⁹ CONSELHO NACIONAL DE JUSTIÇA. Resolução nº 332, de 21 de agosto de 2020.

⁵⁹⁰ Tradução livre de: “40. *Explainability* refers to making intelligible and providing insight into the outcome of AI systems. The explainability of AI systems also refers to the understandability of the input, output and the functioning of each algorithmic building block and how it contributes to the outcome of the systems. Thus, explainability is closely related to transparency, as outcomes and sub-processes leading to outcomes should aim to be understandable and traceable, appropriate to the context. AI actors should commit to ensuring that the algorithms developed are explainable. In the case of AI applications that impact the end user in a way that is not temporary, easily reversible or otherwise low risk, it should be ensured that the meaningful explanation is provided with any decision that resulted in the action taken in order for the outcome to be considered transparent” (UNESCO. Op. cit., p. 9).

⁵⁹¹ Id.

auditoria por autoridade humana quanto a qualquer proposta de decisão apresentada pelo modelo de Inteligência Artificial”⁵⁹², especialmente em relação às de natureza judicial.

Todavia, no plano da Recomendação delineada pela UNESCO, vê-se que, em cada caso, o nível de transparência e explicabilidade deve sempre estar relacionado ao contexto de uso do sistema (aos limites de explicabilidade do próprio sistema de IA em seu uso padrão), pois existe um balanceamento entre a transparência e a explicabilidade e outros princípios, como o da segurança (*safe and security*).⁵⁹³

Nesse sentido, deve-se fazer a distinção entre a **interpretabilidade** (*interpretability*)⁵⁹⁴ do modelo de Inteligência Artificial, a qual se relaciona, no campo da Ciência da Computação, ao grau de compreensão do funcionamento e dos resultados dos processos algorítmicos, o que pode variar a depender da técnica de IA utilizada, e a **explicabilidade** (*explainability*)⁵⁹⁵, que pode ser implementada por meio de diversas técnicas para que os processos e os resultados advindos da Inteligência Artificial expressem explicações em linguagem natural e/ou através de outro meio inteligível pelos seres humanos. Dessa forma, a profunda compreensão dos processos de computação realizados pela IA está no campo da interpretabilidade (a qual nem sempre é alcançável de maneira plena), enquanto a explicabilidade pode ser viabilizada por diversas soluções técnicas a fim de possibilitar a acessibilidade das pessoas afetadas pela decisão. Dessa maneira, um conceito amplo de explicabilidade engloba os conceitos de explicabilidade estrita (destinada às pessoas que não necessariamente tem conhecimento técnico sobre IA) e de interpretabilidade (destinada aos profissionais da Ciência e da Engenharia da Computação, bem como

⁵⁹² Id.

⁵⁹³ Id.

⁵⁹⁴ “*Interpretability: to which extent the model and or the prediction are human understandable. The most addressed discussion is related to how the interpretability can be measured. In a component for measuring the interpretability is the complexity of the predictive model in terms of the model size. According to the literature, we refer to interpretability also with the name comprehensibility*” (GUIDOTTI, Riccardo; MONREALE, Anna; RUGGIERI, Salvatore; TURINI, Franco; GIANNOTTI, Fosca; PEDRESCHI, Dino. A Survey of Methods for Explaining Black Box Models. *ACM Computing Surveys*, v. 51, n. 5, Article 93, January 2019, 42 pages. DOI: 10.1145/3236009).

⁵⁹⁵ “*We take the stance that interpretability alone is insufficient. In order for humans to trust black-box methods, we need explainability – models that are able to summarize the reasons for neural network behavior, gain the trust of users, or produce insights about the causes of their decisions. While interpretability is a substantial first step, these mechanisms need to also be complete, with the capacity to defend their actions, provide relevant responses to questions, and be audited. Although interpretability and explainability have been used interchangeably, we argue there are important reasons to distinguish between them. Explainable models are interpretable by default, but the reverse is not always true*” (GILPIN et al. Op. cit.).

aos profissionais do Direito e de outras ciências sociais que tenham tal conhecimento técnico em IA).

Para além disso, a partir das definições de proporcionalidade, justiça, supervisão humana, transparência e explicabilidade trazidas na Recomendação da UNESCO, e contextualizadas no decorrer deste tópico, pode-se analisar os efeitos das ideias de responsabilidade (*responsibility*) e *accountability*. A **responsabilidade** implica na observância de padrões legais, conforme direitos humanos previstos na legislação internacional e orientações éticas, nas decisões e ações baseadas em sistemas de Inteligência Artificial, sempre atribuindo tal responsabilidade à pessoa, entidade ou organização que desenvolva ou utilize a IA.⁵⁹⁶ Já no que toca à **accountability** devem existir mecanismos adequados de auditabilidade e a rastreabilidade sobre os sistemas de IA e seus resultados, inclusive do ponto de vista dos projetos técnicos e institucionais.⁵⁹⁷

Nesse caminho, a *Policy Area 1* (avaliação de impacto ético) indica expressamente a necessidade do estabelecimento de requisitos claros e objetivos de transparência e explicabilidade para sistemas de Inteligência Artificial voltados para a tomada de decisão na aplicação da lei⁵⁹⁸, inclusive no que se refere ao comportamento de algoritmos responsáveis pela tomada de decisão e aos dados envolvidos, pois a atividade jurisdicional é um de serviço público com necessária interação com o usuário final, conforme expresso na seção 52 da recomendação da UNESCO.⁵⁹⁹

Ademais, de maneira a se garantir, no plano institucional, o incentivo à explicabilidade, também deve ser destacada a *Policy Area 2* (governança e gestão éticas), a qual estabelece que para ser garantida a independência do Judiciário os sistemas de IA de auxílio à decisão judicial devem disponibilizar mecanismos de monitoramento aos órgãos de controle⁶⁰⁰ e fiscalização.⁶⁰¹

Aqui destaca-se que a seção 70 da Recomendação indica que o nível de informação sobre os algoritmos e sobre os resultados de um sistema de IA e a forma

⁵⁹⁶ UNESCO. Op. cit., p. 9.

⁵⁹⁷ Ibid, 9-10.

⁵⁹⁸ Ibid, 9-11.

⁵⁹⁹ “52. Member States and private companies should implement proper measures to monitor all phases of an AI system lifecycle, including the behaviour of algorithms in charge of decision making, the data, as well as AI actors involved in the process, especially in public services and where direct end-user interaction is needed” (Id.).

⁶⁰⁰ No contexto da Recomendação da UNESCO, aqui pode ser entendida também a ideia de controle social.

⁶⁰¹ UNESCO. Op. cit., p. 12.

de explicação necessária podem variar dependendo do público-alvo (*target audience*) que solicita a explicação⁶⁰², como: usuários, especialistas, desenvolvedores, etc.

Por outro lado, também deve ser verificada a viabilidade de cada nível de explicabilidade, pois no atual estado da técnica da Inteligência Artificial e nos limites da tecnologia existente, muitos algoritmos de IA ainda não são totalmente explicáveis, devendo ser observado que para determinados tipos de algoritmos a explicabilidade adiciona um custo de implementação no projeto e na demanda por processamento⁶⁰³. Dessa forma, até que seja tecnicamente possível a explicabilidade completa com o menor impacto na funcionalidade dos modelos de IA, haverá uma troca entre a precisão/qualidade de um sistema de IA e seu nível de explicabilidade⁶⁰⁴. Em suma, um maior nível de explicabilidade impacta negativamente na acurácia da IA, e uma melhor acurácia traz consigo maior grau de opacidade no funcionamento do sistema⁶⁰⁵. O desafio referido foi tratado de maneira incipiente pelo art. 19, *caput*, da Resolução 332/2020 do CNJ ao determinar que “Os sistemas computacionais que utilizem modelos de Inteligência Artificial como ferramenta auxiliar para a elaboração de decisão judicial observarão, como critério preponderante para definir a técnica utilizada, a explicação dos passos que conduziram ao resultado”⁶⁰⁶. O dever de explicar os “passos que conduziram ao resultado” é um começo para orientar a Ciência da Computação na construção de modelos de IA judicial, mas uma IA judicial deve ir além nesse desafio (especialmente nos aspectos inerentes aos testes de racionalidade próprios da teoria da argumentação).

Nesse ponto, reside pesada crítica sobre a aplicação de sistemas de IA à decisão judicial, sendo que na ausência de soluções advindas do próprio método (a Inteligência Artificial em si) e da Ciência da Computação, cabe ao Direito ocupar seu espaço natural para moldar o método a partir da Teoria do Direito e de claros limites jurídicos. Uma abordagem do processamento de linguagem natural (*NLP*) que passe

⁶⁰² Ibid., p. 13.

⁶⁰³ “*xAI is not without costs, however. Most significantly, making an algorithm explainable may result in a decrease in its accuracy. xAI may also stifle innovation, force developers to reveal trade secrets, and impose high monetary costs because xAI can be expensive to build*” (DEEKS, Ashley. Op. cit.)

⁶⁰⁴ “*To avoid this trap, explanations should allow a tradeoff between interpretability and completeness. Rather than providing only simple descriptions, systems should allow for descriptions with higher detail and completeness at the possible cost of interpretability. Explanation methods should not be evaluated on a single point on this tradeoff, but according to how they behave on the curve from maximum interpretability to maximum completeness*” (GILPIN et al. Op. cit.).

⁶⁰⁵ O'NEIL, Cathy. **Weapons of Math Destruction**: how Big Data increases inequality and threatens democracy. New York: Crown, 2016, p. 20. E-book Kindle.

⁶⁰⁶ CONSELHO NACIONAL DE JUSTIÇA. Resolução nº 332, de 21 de agosto de 2020.

pela construção de soluções de justificação racional para as decisões, ainda que não tenha sido pensada ou desenvolvida até o presente momento, é um ponto a ser explorado na interseção entre Teoria do Direito e Ciência da Computação rumo à explicabilidade das decisões tomadas por IA que venham a impactar, direta ou indiretamente, nas decisões judiciais.

Nesse sentido, nos modelos de IA jurídica voltados para a construção de decisões judiciais, o modelo teórico da teoria da argumentação⁶⁰⁷, haja vista sua abordagem analítica e procedimental para a justificação das decisões judiciais⁶⁰⁸, pode servir como base para a construção de modelos de IA que incluam a explicabilidade desde o desenvolvimento do sistema e do processamento de seus resultados⁶⁰⁹ ou que, mediante um sistema auxiliar, criem a possibilidade de explicabilidade para os resultados obtidos mediante *machine learning* ou *deep learning*.

Nesse caminho, a seção 68 da Recomendação da UNESCO demanda a adaptação e a revisão dos marcos regulatórios e legais do país, de maneira a garantir a responsabilização, pelo conteúdo e pelos resultados, nas diversas fases do ciclo de vida da IA, bem como o dever de *accountability* de pessoas físicas e jurídicas responsáveis.⁶¹⁰

Desse modo, deve existir a avaliação de impacto da implementação do sistema de IA (seção 50 da Recomendação da UNESCO), mediante procedimento democrático (com possibilidade de participação dos cidadãos), na qual deve ser realizada a avaliação de benefícios e riscos (aos direitos humanos, ao meio ambiente, à ética e à sociedade), além de medidas de prevenção, mitigação e monitoramento

⁶⁰⁷ ATIENZA, Manuel. Op. cit., p. 192-219.

⁶⁰⁸ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 261-264.

⁶⁰⁹ “In Phase 2 (the coding phase), the AI is developed and worked on to ensure it meets the necessary standards. This is often a point of introduction for problems related to bugs and biases, and is well described in the literature (e.g., Sweeney 2013; Datta et al. 2015; O’Neil 2016; Boscoe 2019). In this phase, it is discussed what the accuracy rates are, what they should be, how these and other performance metrics differ, how they should be allowed to differ across different subpopulations (when deciding about groups or individuals), what data to use when training the algorithm, and how to clean it. With public decision-making, the main challenge is ensuring that the AIs are good enough when it comes to these issues. Of course, it may be difficult or even impossible for decision-makers to know for sure whether the AIs they have authorized are up to standard without relying on programmers. This is not a problem restricted to AI decision-making, since decision-makers rely on expert opinions in virtually all policy areas. However, the problem may be accentuated in AI decision-making, as few political representatives are trained in code reading or programming. Furthermore, it may be difficult to establish goals and guarantee that their respective importance is sufficiently precise for programmers’ needs” (LICHT, K. de Fine, LICHT, J. de Fine. Op. cit.).

⁶¹⁰ UNESCO. Op. cit. p. 13.

de riscos⁶¹¹. Para isso o setor público (no que se inclui o Poder Judiciário) deve realizar autoavaliação dos sistemas de IA existentes ou propostos, no que toca a adequação do uso da IA e os métodos utilizados na implementação desta. Para tanto, conforme se verifica na seção 51 da Recomendação da UNESCO, devem ser estabelecidos mecanismos de supervisão adequados (para auditabilidade, rastreabilidade e explicabilidade)⁶¹².

A avaliação dos sistemas públicos de IA deve ser multidisciplinar⁶¹³, envolvendo as diversas partes interessadas, como atores do ecossistema de IA, representantes da sociedade civil, servidores públicos, seguradoras, investidores, fabricantes, engenheiros, advogados e usuários (seção 69 da Recomendação da UNESCO)⁶¹⁴, bem como deve ser multicultural, pluralista e inclusiva, o que pode incluir uma fase de revisão externa (controle social).⁶¹⁵ Tal esforço, especialmente para a auditabilidade e a rastreabilidade, serve para que eventuais danos causados aos usuários (jurisdicionados) através de sistemas de IA possam ser investigados, punidos e reparados.

Já a seção 77 da Recomendação da UNESCO traz o dever de abordagem dos dados voltada para a interoperabilidade dos bancos de dados vinculados aos sistemas do poder público⁶¹⁶, a fim de possibilitar o compartilhamento de base de dados de qualidade em um espaço comum ao poder público, o qual deve ser confiável e protegido. Percebe-se que tal dever de interoperabilidade ainda é um desafio não vencido pelo Poder Judiciário do Brasil, especialmente em relação ao art. 194 do Código de Processo Civil⁶¹⁷ e ao art. 24, I, da Resolução 332/2020 do Conselho

⁶¹¹ Ibid., 10-11.

⁶¹² Ibid., p. 11.

⁶¹³ *“On top of this smaller group of genuine legal specialists and this growing cadre of hybrids, I also envisage the emergence of a further grouping of professionals — the legal knowledge engineers. These are the highly skilled individuals who will be engaged in the jobs of standardizing, systematizing, and packaging the law. They will be the analysts who reorganize and restructure legal knowledge in a form that can be embodied in advanced systems, whether for use by lawyers, paralegals, or lay people”* (SUSSKIND, Richard. **The End of Lawyers?**, p. 7).

⁶¹⁴ UNESCO. Op. cit., p. 13.

⁶¹⁵ Ibid., p. 11.

⁶¹⁶ Ibid., p. 14.

⁶¹⁷ “Art. 194. Os sistemas de automação processual respeitarão a publicidade dos atos, o acesso e a participação das partes e de seus procuradores, inclusive nas audiências e sessões de julgamento, observadas as garantias da disponibilidade, independência da plataforma computacional, acessibilidade e interoperabilidade dos sistemas, serviços, dados e informações que o Poder Judiciário administre no exercício de suas funções”.

Nacional de Justiça.⁶¹⁸

O panorama até aqui descrito indica a criação de um ambiente normativo e institucional que estimula a explicabilidade como o elemento central de coesão entre os princípios de confiança, responsabilidade, *accountability* e justiça da Inteligência Artificial jurídica.

2.7 IA Explicável (xAI) e Transparência Possível: Os limites da Explicabilidade da Inteligência Artificial nas Decisões Judiciais

A opacidade das decisões tomadas por meio a IA (especialmente via *machine learning*) lança um desafio à democracia, pois nesta, por definição, as decisões são tomadas de maneira que a transparência e a publicidade garantam⁶¹⁹, mais que o acesso à própria decisão, o questionamento dos motivos das decisões.

O modelo democrático de controle dos atos estatais ganha traços ainda mais sérios no campo da decisão judicial, pois para a garantia do Estado de Direito há a necessidade de explicabilidade dos motivos pelos quais uma decisão judicial é tomada.

Desse modo, a explicabilidade estrita de uma decisão judicial ocorre através da motivação dessa, a qual, para a teoria do Direito, se dá através da adequada justificação racional pela via das teorias da argumentação. Justificar uma decisão é dar uma explicação lógico-dialética no plano argumentativo (sem que isso implique na exclusão dos elementos retóricos ou poéticos das decisões).

Contudo, a transparência dos algoritmos utilizados em decisões auxiliadas por IA pode, a depender do desenvolvimento de interface adequada entre o Direito e o modelo tecnológico, ser mais facilmente obtida de sistemas puramente produzidos

⁶¹⁸ “Art. 24. Os modelos de Inteligência Artificial utilizarão preferencialmente software de código aberto que:

I – facilite sua integração ou interoperabilidade entre os sistemas utilizados pelos órgãos do Poder Judiciário” (CONSELHO NACIONAL DE JUSTIÇA. Resolução nº 332, de 21 de agosto de 2020).

⁶¹⁹ “One of the early progressives’ basic insights about transparency ought to resonate across the political spectrum. In general, they believed, the case for exposure is strongest where secrecy enables the accumulation of arbitrary political or economic power”. (...) “The risk of undisclosed and unchecked domination in the United States is today sourced less plausibly to legislatures or political machines, and more plausibly to institutions such as multinational corporations or the defense, intelligence, and law enforcement agencies — and to the computer algorithms used by all of these. Wherever the potential for such domination resides, transparency reforms that seek to reduce it ought to be a higher priority than reforms that seek marginally more openness from organizations that are already highly visible” (POZEN, David E. Transparency’s Ideological Drift. *The Yale Law Journal*, v. 128, n.1, p. 100-165, 2018).

pelo poder público, pois os sistemas do Poder Judiciário devem usar, preferencialmente, programas com código aberto (art. 14, *caput*, da Lei 11.419/2006 e art. 24, *caput*, da Res. 332/2020 do CNJ⁶²⁰).

Por outro lado, os programas de computador desenvolvidos por agentes privados são considerados segredos industriais pela legislação⁶²¹. Assim, é necessária cautela na contratação pelo Poder Judiciário (por meio de contrato de licença de *software*)⁶²² do fornecimento Inteligência Artificial por entidades privadas, de maneira a garantir a plena auditabilidade dos algoritmos e dos dados.

No âmbito das possibilidades de implementação da IA, a criação de sistemas de Inteligência Artificial voltados à decisão judicial deve ser integralmente estruturada pelos Tribunais (art. 25 da Res. 332/2020 do CNJ⁶²³), ainda que em colaboração com entidade pública ou privada de pesquisa científica⁶²⁴, uma vez que o ato de julgar caracteriza exercício de função típica e indelegável do Poder Judiciário (art. 2º, *caput* e art. 5º, XXXV, da Constituição Federal).

Portanto, é vedada a delegação do processamento (no todo ou em parte) necessário às decisões judiciais para serviços de *cloud computing* de entidades privadas (via API⁶²⁵ do *IBM Watson*⁶²⁶ ou do *GPT-4* da *Open AI*⁶²⁷, por exemplo) fora do alcance direto do Poder Judiciário, haja vista a opacidade⁶²⁸ ocasionada pela dificuldade de acesso ao processamento realizado fora dos servidores de cada Tribunal, o que caracterizaria (ou potencializaria) um modelo *black box*⁶²⁹ de difícil auditabilidade e baixa explicabilidade.

⁶²⁰ CONSELHO NACIONAL DE JUSTIÇA. Resolução nº 332, de 21 de agosto de 2020.

⁶²¹ Art. 20, §1º, da Lei 13.709/2020, art. 5º, XXVII, da Constituição Federal, art. 2º, *caput* e §5º, art. 3º, §1º, III, e §2º, da Lei 9.609/1998 e Lei 9.610/1998.

⁶²² Na forma do art. 7º da Lei 9.609/1998 e que envolve inexigibilidade de licitação (art. 25, I, da Lei 8.666/1993 e art. 72, I, da Lei 14.133/2021).

⁶²³ CONSELHO NACIONAL DE JUSTIÇA. Resolução nº 332, de 21 de agosto de 2020.

⁶²⁴ Art. 219-A da Constituição Federal e art. 28 da Res. 332/2020 do CNJ.

⁶²⁵ API é a sigla para *Application Programming Interface*, o que representa o acesso a serviço de computação disponibilizado por terceiro via Internet no local em que se encontra o usuário. Nesse sentido: GREENFIELD, Adam. Op. cit., p. 83.

⁶²⁶ IBM. **Build with Watson: The AI for business**. [2020]. Disponível em: <https://www.ibm.com/watson/developer>. Acesso em: 09 ago. 2020.

⁶²⁷ OPENAI. Op. cit.

⁶²⁸ “The inability to parse the reasons behind the algorithm’s recommendations can harm those affected by the recommendations. Opaque algorithms can undercut people’s sense of fairness and trust – particularly when used by the government – and in the criminal justice setting can undercut a defendant’s right to present a defense” (DEEKS, Ashley. Op. cit.).

⁶²⁹ “Yet in a climate of secrecy, bad information is as likely to endure as good, and to result in unfair and even disastrous predictions. This is why the wholesale use of black box modeling, however profitable it is for the insiders who manage it, is dangerous to society as a whole. It’s bad enough when innocent individuals are hurt, branded as security threats or goldbrickers or credit risks by inaccuracies that they

O baixo grau de transparência não pode ser aceito no âmbito da construção de decisões judiciais com o auxílio de Inteligência Artificial. A manutenção de IA judicial (como nos casos de processamento de linguagem natural, por exemplo) deve ser feita no ambiente do próprio Judiciário, a fim de existir controle sobre eventuais desvios e vieses executados no sistema computacional, tendo em conta que o julgamento é tarefa exclusiva do Poder Judiciário. Abre-se mão do poder computacional entregue via a API de uma *big tech*, mas se garante níveis mínimos de explicabilidade e a transparência inerentes aos padrões éticos. Uma hipótese de possível aceitação do modelo decorreria de disponibilização dos algoritmos e dos processos da detentora da IA encapsulada para auditoria externa, o que se mostra difícil, haja vista a proteção ao segredo de negócio que alcança tais empresas. Demarcados os referidos limites, cabem outros pontos a serem observados.

Destaca-se, no atual estado da técnica da Ciência da Computação, que não há resposta plena para o anseio de parcela da comunidade jurídica por um tipo de explicabilidade completa (*full explainability*), onde cada *bit* de processamento teria uma causalidade lógica que poderia ser traduzida numa espécie de mapa mental acessível ao operador do Direito e a qualquer cidadão. A explicabilidade completa (explicabilidade em sentido amplo) é um desafio que ainda não pode ser alcançado, haja vista a) que a interpretabilidade (*interpretability*) dos modelos de IA é variável e, muitas vezes, não é adequadamente alcançada nem mesmo por cientistas da computação e engenheiros de *machine learning*; e b) que existe uma troca (*trade-off*) entre maior explicabilidade de sistemas com menor acurácia ou melhores acurácia e sensibilidade de modelos de mais difícil explicabilidade, havendo nesse caso uma explicabilidade em grau menor, que pode ser denominada de explicabilidade em sentido estrito (a qual envolve a explicabilidade do modelo algorítmico desenvolvido sem que haja a plena interpretabilidade dos resultados e processos computacionais).

Por outro lado, a partir da referida constatação, uma pergunta deve ser feita (e respondida) para que a premissa mais básica acerca da explicabilidade de decisões judiciais seja estabelecida: quais são os tipos e os graus de explicabilidade exigidos dos juízes humanos?

can't contest and may not even know about. Modeling is even worse when unfair or inappropriate considerations combine with the power of algorithms to create the failures they claim to merely predict" (PASQUALE, Frank. **The Black Box Society**, p. 216-217).

Para problematizar a pergunta, percebe-se que, na teoria do Direito, há uma demanda pela objetivação da decisão judicial, na medida em que a explicabilidade da mesma passa pela exigência de: a) que a decisão seja fundamentada (art. 93, IX, da Constituição Federal); b) que a fundamentação seja analítica (art. 489, §1º, do Código de Processo Civil); c) que a decisão tenha se baseado em fundamentos lógico-rationais próprios da razão prática⁶³⁰; d) que a decisão tenha se baseado em fundamento(s) jurídico(s)⁶³¹; e) que tenha sido aplicado corretamente o Direito⁶³²; f) que nas hipóteses admitidas (pseudolacunas do sistema⁶³³ ou abertura da linguagem⁶³⁴) ou estimuladas (adaptabilidade dos princípios⁶³⁵) haja a possibilidade de se criar direito novo (que poderá⁶³⁶ ser um novo precedente⁶³⁷); e g) que haja caráter institucional⁶³⁸ que indique a legitimidade da decisão. Logo, as bases da explicabilidade da decisão judicial, ainda que exista debate no plano teórico sobre aspectos específicos, têm elementos centrais já bastante sedimentados.

Nesse sentido, a percepção da motivação interna, psicológica, de por que um juiz proferiu determinada decisão acaba não interessando ao Direito, uma vez que a motivação interna para a tomada de uma decisão causa influência apenas acidental (hipóteses de impedimento e suspeição do julgador⁶³⁹).

A opacidade (e a aleatoriedade) do cérebro humano é muitíssimo maior⁶⁴⁰ do que se verifica nos modelos de Inteligência Artificial. Dessa forma, na normalidade dos casos, não se pode, por exemplo, pedir uma análise psicológica ou psiquiátrica

⁶³⁰ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 189.

⁶³¹ Ibid., p. 195-196.

⁶³² Ibid., p. 189-192.

⁶³³ KELSEN, Hans. **Teoria Geral do Direito e do Estado**, p. 215.

⁶³⁴ HART, Herbert. Op. cit., p. 164.

⁶³⁵ ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 90.

⁶³⁶ A palavra “poderá” denota exatamente o desafio que é manter uma nova interpretação através do filtro institucional de formação de precedentes nos Tribunais locais e nos Tribunais superiores.

⁶³⁷ BARBOZA, Estefânia Maria de Queiroz. Op. cit., p. 22-23.

⁶³⁸ “Ao (antigo) questionamento sobre quem possui as melhores condições para responder a casos difíceis, antepõe-se a resposta de que a construção dos saberes jurídicos precisa ser realizada como uma conversa contínua entre todos aqueles que se encontram envolvidos em sua produção e revisão” (LORENZETTO, Bruno Meneses; CLÈVE, Clèmerson Merlin. *Teorias Interpretativas, Capacidades Institucionais e Crítica*).

⁶³⁹ Art. 144 e art. 145 do Código de Processo Civil.

⁶⁴⁰ “Apesar de haver uma grande quantidade de processos estocásticos (aleatórios dentro de restrições cuidadosamente controladas) em todos os aspectos do cérebro, não é necessário modelar toda ‘cavinha’ na superfície de todo dendrito, mais do que é necessário modelar todas as pequenas variações na superfície de todos os transistores, para compreender os princípios operacionais de um computador. Mas certos detalhes são críticos para decodificar os princípios operacionais do cérebro, o que nos força a distinguir entre eles e aqueles que compõem ‘ruído’ estocástico ou caos” (KURZWEIL, Ray. **A Singularidade Está Próxima**, p. 174).

da pessoa do julgador para tentar compreender os motivos que o levaram a uma decisão. Haveria maior grau de explicabilidade com a submissão do cérebro do julgador à análise de profissionais da saúde mental mediante abordagens da psicologia (conforme os estudos desenvolvidos por Daniel Kahneman⁶⁴¹ ou por Paul Ekman⁶⁴², por exemplo) e da neurociência (via exames de imagem por ressonância magnética funcional – fMRI⁶⁴³), o que, todavia, são métodos invasivos vedados. Ademais, a tentativa de explicabilidade dos pensamentos que levaram à decisão judicial seria inócua, pois, ainda que um juiz discorde de sua sentença ou que tenha sido movido por motivos extrajurídicos, se “ele interpretou o Direito de forma correta ou incorreta, deduz-se da sua sentença”⁶⁴⁴ de maneira objetiva. O exemplo torna claro o ponto em debate. Todavia, destaca-se, aqui, que a auditabilidade deve sempre ser buscada, mas mesmo para essa existem limites técnicos, essa é uma realidade que irá nortear as escolhas no processo de desenvolvimento dos modelos de IA judicial.

Já para a Inteligência Artificial existem abordagens, arquiteturas e ferramentas que podem, em variados graus, fazer enxergar além do biombo da opacidade. Desse modo, a opacidade psicológica torna insondável o processamento mental do juiz, já o tratamento da opacidade da Inteligência Artificial jurídica, para além do ato de se arranhar a superfície de resultados computacionais, é uma questão de técnica, a qual em grande parte já vem sendo desenvolvida e avançará ainda mais.

⁶⁴¹ “Os participantes inadvertidos do estudo eram oito juízes de condicional em Israel. Eles passam dias inteiros revisando pedidos de condicional”. (...) “Os autores do estudo fizeram um gráfico da proporção de pedidos aprovados em relação ao tempo desde a última pausa para refeição. A proporção conhece picos após cada refeição, quando cerca de 65% dos pedidos são concedidos. Durante as duas horas, mais ou menos, até a refeição seguinte dos juízes, a taxa de aprovação cai regularmente, até chegar perto de zero pouco antes da refeição. Como era de se esperar, esse é um resultado indesejável e os autores verificaram cuidadosamente muitas explicações alternativas. A melhor explicação possível dos dados é uma má notícia: juízes cansados e com fome tendem a incorrer na mais fácil posição *default* de negar os pedidos de condicional. Tanto o cansaço como a fome provavelmente desempenham um papel”. (KAHNEMAN, Daniel. Op. cit., posição 897-902).

⁶⁴² “*Much of the time, for some people all of the time, our emotions serve us well, by mobilizing us to deal with what is most important in life and providing us with many different kinds of enjoyment. But sometimes our emotions get us into trouble*” (EKMAN, Paul. **Emotions Revealed: Understanding Faces and Feelings**. London: Orion, p. 30. E-book Kindle).

⁶⁴³ NISHIMOTO, Shinji; VU, An T.; NASELARIS, Thomas; BENJAMINI, Yuval; YU, Bin; GALLANT, Jack L. Reconstructing Visual Experiences from Brain Activity Evoked by Natural Movies. **Current Biology**, v. 21, p. 1641-1646, Berkeley, 2011. DOI: 10.1016/j.cub.2011.08.031.

⁶⁴⁴ BULYGIN, Eugenio. Alexy e o Argumento da Correção. In: BULYGIN, Eugenio; ALEXY, Robert. **Direito, Moralidade e Positivismo: o debate entre Robert Alexy e Eugenio Bulygin**. Curitiba: Juruá, 2017, p. 77.

Nesse sentido, o sistema jurídico deve formular pretensão de correção⁶⁴⁵ e satisfazer tal pretensão, motivo pelo qual a opacidade de decisões tomadas por sistemas de Inteligência Artificial jurídica deve ser compreendida nos limites e possibilidades do estado da técnica da Ciência da Computação, limites esses que devem ser permeados pelo Direito, sabendo que a explicabilidade é um problema que não nasceu com a IA, mas com o próprio ser humano.

Para além do direito fundamental à proteção de dados⁶⁴⁶ previsto no art. 5º, LXXIX, da Constituição Federal, a transparência algorítmica e a explicabilidade podem ser alçadas à condição de direitos fundamentais atribuídos⁶⁴⁷, em decorrência do alcance do bloco normativo-constitucional de garantia de acesso às informações, de publicidade e de *accountability* das decisões do poder público, especialmente do Poder Judiciário.

Destaca-se, mais uma vez, que a explicabilidade é alcançada em diversos graus. Mesmo nos casos em que não haja explicabilidade completa (*full explainability*), advinda de plena interpretabilidade, há a possibilidade de implementação de explicabilidade do processo algorítmico mediante sistema sobreposto⁶⁴⁸ que crie, selecione ou classifique a argumentação jurídica para o **teste de juridicidade** dos resultados (*outputs*) trazidos pela Inteligência Artificial. Observa-se que o referido **teste de juridicidade** funciona como um mecanismo de emergência⁶⁴⁹ apto a denunciar decisões advindas de vieses e outros defeitos⁶⁵⁰.

⁶⁴⁵ “A pretensão de correção que se coloca nos discursos jurídicos se distingue claramente da do discurso prático geral. Não se pretende que o enunciado jurídico normativo afirmado, proposto ou ditado como sentença seja só racional, mas também que no contexto de um ordenamento jurídico vigente possa ser racionalmente fundamentado. Isso se refere exatamente ao que deve ser esclarecido no âmbito de uma teoria do discurso jurídico” (ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 189-190).

⁶⁴⁶ DONEDA, Danilo. A proteção dos dados pessoais como um direito fundamental. **Espaço Jurídico Journal of Law**, v. 12, n. 2, p. 91-108, 13 dez. 2011.

⁶⁴⁷ “Uma atribuição correta ocorre quando a norma atribuída pode ser classificada como válida. Para classificar as normas diretamente estabelecidas pelo texto constitucional basta a referência à sua positividade” (ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 74).

⁶⁴⁸ “An alternative to these exogenous approaches is a category of xAI that attempts to explain (or “decompose”) the model’s reasoning. The most obvious way to do so is to reveal the source code for the machine learning model, but that approach will often prove unsatisfactory (because of the way machine learning works and because most people will not be able to understand the code). More nuanced alternatives exist, however. One approach is to create a second system alongside the original ‘black box’ model, sometimes called a ‘surrogate model.’ A surrogate model works by analyzing featured input and output pairs but does not have access to the internal weights of the model itself” (DEEKS, Ashley. Op. cit.).

⁶⁴⁹ BOSTROM, Nick. Op. cit., posição 4780-4809.

⁶⁵⁰ “Os processos internos da IA poderiam ser monitorados à procura de atividades suspeitas. Por exemplo, qualquer mudança na representação dos valores finais da IA poderia resultar em um desligamento automático para uma revisão. A maneira como a IA utiliza recursos e métodos

Nesse caminho, a dualidade de sistemas é estruturada de maneira que o sistema decisório primário encontre as decisões que entenda como passíveis de constituírem proposições jurídico-normativas, e que o sistema secundário realize o teste de juridicidade sobre o resultado fornecido.

Em suma, no desafio da transparência algorítmica da IA, os resultados: a) podem ser desvelados em graus variados a depender da maneira de construção pelos desenvolvedores; b) podem permanecer com seu processamento perdido na caixa preta (fator *black box*); c) podem ser explicados com uma IA desenhada para referida tarefa; ou d) podem ser limitados por meio do teste de juridicidade submetido por uma IA verificadora.

Nesse sentido, a interpretabilidade e a explicabilidade das decisões não são problemas novos, mas são questões tradicionais acerca da transparência e da fundamentação do que é decidido pelo Poder Judiciário.

O direito fundamental à explicabilidade das decisões judiciais tomadas com auxílio de IA ficou caracterizado, desde o bloco de normas constitucionais e legais que podem ser aplicadas às ideias de *accountability* e de proteção de dados, passando por uma visão otimizada do art. 20 da Lei Geral de Proteção de Dados Pessoais até à construção dos conceitos de interpretabilidade e explicabilidade a partir de diversas iniciativas no plano internacional, como os Princípios de Asilomar, a Recomendação 449 da OCDE e a Recomendação da UNESCO sobre ética em IA. A Resolução 332/2020 do CNJ⁶⁵¹ também é instrumento normativo que buscou tratar da explicabilidade das decisões judiciais tomadas com o auxílio de IA.

A explicabilidade não é, portanto, um impeditivo ao desenvolvimento científico (art. 5º, IX, e art. 218, *caput* e §1º, da Constituição Federal) da Inteligência Artificial jurídica, mas sim uma força tecnológica fruto do progresso da ciência (valor protegido no plano constitucional) e que tem limites próprios de seu estado da arte.

Por outro lado, a precaução necessária na pesquisa e no desenvolvimento de soluções pautadas na Inteligência Artificial para o Judiciário passa pelo necessário respeito aos valores éticos. Não se trata de bloquear o desenvolvimento e a pesquisa em Inteligência Artificial, o que, do ponto de vista puramente normativo e regulatório

algorítmicos poderia também ser monitorada contra padrões inesperados. De modo mais ambicioso, os pontos de vista da IA, planos, submetas e justificativas, poderiam ser continuamente inspecionados para assegurar que estejam de acordo com as expectativas dos programadores" (Ibid., posição 4826-4843).

⁶⁵¹ CONSELHO NACIONAL DE JUSTIÇA. Resolução nº 332, de 21 de agosto de 2020.

soaria irreal, uma vez que a evolução tecnológica não costuma respeitar barreiras puramente normativas.

O fenômeno da regulação da Inteligência Artificial tem abrangência internacional, fugindo ao controle e às fronteiras dos estados. Assim, é necessário o reconhecimento do direito à transparência das decisões tomadas com base em Inteligência Artificial, de maneira a ser tratada no plano interno do direito positivo (e das práticas) dos estados, na condição de um direito fundamental, e no plano internacional e supranacional, através de direito humano específico.

Evidencia-se que a nova camada de linguagem da decisão judicial, implementada através da Inteligência Artificial, traz consigo o risco de opacidade, que deve ser tratado na computação jurídica mediante a criação de soluções híbridas de mútua cooperação entre Teoria do Direito e Ciência da Computação, considerando que já existem arquiteturas de sistemas voltados para a explicabilidade de decisões tomadas por IA. Deve-se ainda destacar que o interprete humano deve estar inserido na condição de supervisor e validador último dos resultados obtidos pela IA, pois apenas o Juiz (humano) é o legítimo agente no sistema constitucional para a tomada da decisão.

Nesse aspecto, a teoria da justificação racional das decisões judiciais tem elementos que podem ser utilizados para potencializar as técnicas já existentes de Inteligência Artificial explicável (xAI), sendo a teoria da justificação a base para a construção de sistemas de IA jurídica adequadamente explicáveis.

Por outro lado, a definição dos conceitos de transparência e explicabilidade (num sentido amplo) é necessária, especialmente porque a explicabilidade em sentido amplo pode ser subdividida em **interpretabilidade**, que se relaciona aos processos computacionais da IA, aos seus resultados e à sua compreensão pelos profissionais da tecnologia, e **explicabilidade em sentido estrito**, voltada à criação de modelos em linguagem natural ou outros tipos de representação que sejam inteligíveis aos seres humanos destinatários da decisão.

Dessa forma, a dinamicidade, a adaptabilidade, a fluidez e a reatividade da linguagem jurídica, ainda que sejam problemas computacionais complexos, vêm sendo tratadas por sistemas de Inteligência Artificial, o que demanda uma arquitetura de decisões com dualidade de implementação algorítmica, uma no plano da tomada de decisões judiciais com auxílio da IA (o que já vem sendo feito pelos Tribunais com

grau de eficiência) e outra no plano da criação de sistemas, sobrepostos ou em paralelo, que executem o teste de juridicidade em relação às decisões tomadas. Essa última cria um desafio justificatório inclusive ao magistrado (humano), que tem sua carga de justificação aumentada quando venha a adotar decisão que não passe despercebida pelo teste de juridicidade.

Sendo assim, inclusive para fins de desenvolvimento, controle e supervisão de sistemas de IA jurídica, é um caminho natural que os profissionais do Direito do futuro (e do presente) passem da figura do especialista (comumente chamado operador do direito) para posições mais próximas do que se pode chamar de desenvolvedor de *software* jurídico e de jurista da computação jurídica.

Portanto, a IA jurídica e a interpretação do direito por meio de pensamento/processamento híbrido humano-máquina já ocorre, e o direito à explicabilidade é passível de reconhecimento com os diplomas normativos já existentes.

Sendo assim, a partir dos estudos desenvolvidos observa-se: a) a necessidade de desenvolvimento de modelos de Inteligência Artificial judicial com interpretabilidade e explicabilidade advindas de transparência *by design* em respeito aos limites éticos e normativos necessários para a integração da IA à decisão judicial; b) o reconhecimento da existência de limites técnicos no direito à explicabilidade das decisões judiciais tomadas com auxílio de Inteligência Artificial; e c) a necessidade de inclusão de critérios jurídicos (mediante teste de juridicidade) nos modelos de explicabilidade para os sistemas de Inteligência Artificial judicial.

Isso será alcançado mediante a necessária interseção entre a teoria da justificação da decisão judicial e as técnicas de Inteligência Artificial explicável (xAI), a fim de garantir a *accountability* das decisões judiciais mediante teste de juridicidade dos resultados fornecidos por sistemas de IA judicial; e a definição do direito fundamental à explicabilidade estrita da decisão judicial tomada com auxílio de IA, de maneira de diferenciá-lo da explicabilidade em sentido amplo (que engloba também a interpretabilidade dos processos computacionais necessários ao funcionamento da IA).

Nesse sentido, se, por um lado, a explicabilidade é necessária para o adequado controle dos sistemas de Inteligência Artificial judicial, por outro, deverá ocorrer *accountability* periódica, o que prevenirá problemas nos sistemas mediante a garantia

de que haja uma alimentação ininterrupta da base de dados com novas decisões e novos argumentos passíveis de utilização para cada tipo de caso no escopo de cada *software* de IA judicial.

A partir dessa abordagem, a explicabilidade é o elemento que auxilia no processo decisório em relação aos seguintes aspectos, aqui apresentados de forma sintética e para fins didáticos:

I - “Ler” e “compreender” (carregamento, preparação e extração dos dados) o contexto do processo e dos argumentos utilizados mediante:

- a) Análise do processo como um todo;
- b) Análise de determinados atos do processo;

II - Definição da proposta de decisão:

- a) Análise estatística pura, via técnicas de aprendizado de máquina (*machine learning*);
- b) Análise por meio de programação procedural;
- c) Análise por meio de programação orientada à objetos;
- d) Análise mista em relação qualquer das técnicas anteriores.

A partir desses aspectos, para a computação jurídica, o desenho institucional de sistema de Inteligência Artificial de auxílio à decisão judicial deve levar em consideração as hipóteses de **decisão aplicada pela IA**, que consiste na possibilidade de selecionar sugestões de decisão entre as opções apresentadas pelo sistema, conforme banco de decisões inerentes a uma estruturação de raciocínio baseado em casos (*case-based reasoning*).

Por outro lado, para que haja a adequada explicabilidade também deve ser considerada a implementação de sistemas cujo paradigma de funcionamento seja a **decisão fundamentada pela IA**, no qual a lógica⁶⁵² (proposicional e/ou de

⁶⁵² “Lógicas de argumentação derrotável encontraram ampla aplicação em representação de conhecimento jurídico, a partir de duas abordagens principais: o raciocínio a partir de precedentes (*case based reasoning*) e o raciocínio a partir de regras (*rule based reasoning*). Essas lógicas modelam o raciocínio como inferências a partir de argumentos a favor ou contra determinada tese, incluindo informação sobre a força relativa desses argumentos. Um argumento pode ser uma estrutura inferencial complexa, que liga suas premissas a conclusões por meio de passos intermediários detalhados. Esses argumentos, por sua vez, podem ser atacados em diferentes junções e de diversas formas, e uma

proposições normativas) se volta: a) à estruturação dos sistemas decisórios; b) à estruturação dos microssistemas; e c) à possibilidade de ser utilizada na programação de sistemas baseados em conhecimento jurídico.

Dessa maneira, percebe-se que existem níveis diferentes, com abordagens diversas, as quais podem atuar de maneira isolada ou de maneira mista na busca de construção de sistemas de auxílio à decisão judicial.

Em suma, não existe uma única forma de programar sistemas de decisão judicial baseados em Inteligência Artificial. Todavia, deve ser respeitada a premissa de explicabilidade pautada na construção de um processo argumentativo dinâmico e típico do duplo canal de correção para a interação argumentativa Juiz-IA. Por isso, pensar a explicabilidade no nível da teoria do Direito envolve o esforço para se estabelecer uma teoria de explicabilidade que trate as linhas gerais do projeto, do desenvolvimento (programação), da implementação e da manutenção dos sistemas de IA judicial, segundo seus aspectos jurídicos.

2.8 Aspectos Dogmáticos do Direito à Explicabilidade

Neste ponto da obra, destaca-se que todo o debate sobre o direito à explicabilidade foi aqui contextualizado devido ao fato de constituir uma das bases jurídico-epistêmicas para a própria existência do duplo canal de correção na aplicação do direito com o auxílio de sistemas de Inteligência Artificial.

Conforme referido em tópicos anteriores, este estudo trouxe esforços inerentes ao delineamento do direito à explicabilidade através do art. 20 da Lei Geral de Proteção de Dados (Lei 13.709/2018)⁶⁵³ e de sua leitura ampliada, por meio da força normativa do direito fundamental à proteção de dados pessoais e do princípio da publicidade das decisões judiciais (art. 5º, LXXIX, e art. 93, IX, da Constituição Federal – vide subcapítulos 2.1 e 2.5, em especial).

A construção normativa do direito à explicabilidade de decisões tomadas ou auxiliadas por Inteligência Artificial tem ocorrido em várias frentes a partir de

conclusão pode ser derivada se for possível construir um argumento a favor da conclusão que seja defensável contra todos os argumentos que o atacam” (MARANHÃO, Juliano Souza de Albuquerque; FLORENCIO, Juliana Abrusio; ALMADA, Marco. Inteligência artificial aplicada ao direito e direito da inteligência artificial. **SUPREMA**: Revista de Estudos Constitucionais, Brasília, v. 1, n. 1, p. 154-180, jan/jun. 2021).

⁶⁵³ BRASIL. Lei 13.709, de 14 de agosto de 2018.

preocupações com os riscos surgidos na inclusão de IA no processo de decisão humana (conforme relatado no subcapítulo 1.4).

Além disso, conforme já referido no subcapítulo 2.5, a normatividade do direito à explicabilidade também foi explorada em relação à Recomendação 449 da OCDE⁶⁵⁴, à *Recommendation on the Ethics of Artificial Intelligence*⁶⁵⁵ da UNESCO e à Resolução 332/2020 do Conselho Nacional de Justiça.⁶⁵⁶

Desse modo, tendo em vista as incursões já realizadas em face dos aspectos dogmáticos do direito à explicabilidade no que se refere aos diplomas normativos referidos (Recomendações da OCDE e da UNESCO, Constituição Federal, Lei Geral de Proteção de Dados e Resolução 332/2020 do Conselho Nacional de Justiça), bem como levando em conta que neste Capítulo 2 foram lançadas as bases epistêmicas para compreensão da explicabilidade como um pressuposto ao estudo de sua normatividade, passa-se a uma análise dos demais aspectos dogmáticos envolvidos. Nesse rumo, será feita uma análise dos demais diplomas normativos aplicáveis e de quais são seus efeitos na definição do direito à explicabilidade.

Inicialmente, e com os cuidados necessários ao estudo de direito comparado e o respeito ao contexto cultural de cada agente normativo pelo mundo⁶⁵⁷, assim como tendo em vista as diversas espécies de relações políticas e comerciais do Brasil com países europeus, é necessário destacar as diretrizes surgidas naquele contexto supranacional. Dessa maneira, nas *Ethical Guidelines for Trustworthy AI*⁶⁵⁸ (Diretrizes Éticas para IA Confiável) da União Europeia há a busca pelo estabelecimento de uma concepção de Inteligência Artificial confiável através de três componentes a serem respeitados durante todo o ciclo de vida do sistema: a) **licitude**, mediante cumprindo todas as leis e regulamentos aplicáveis⁶⁵⁹; b) **ética**, através da adesão a princípios e

⁶⁵⁴ OCDE. Op. cit.

⁶⁵⁵ UNESCO. Op. cit.

⁶⁵⁶ CONSELHO NACIONAL DE JUSTIÇA. Resolução 332, de 21 de agosto de 2020.

⁶⁵⁷ LEGRAND, Pierre. Op. cit., posição 223.

⁶⁵⁸ UNIÃO EUROPEIA. **Ethics Guidelines for Trustworthy AI**.

⁶⁵⁹ “AI systems do not operate in a lawless world. A number of legally binding rules at European, national and international level already apply or are relevant to the development, deployment and use of AI systems today. Legal sources include, but are not limited to: EU primary law (the Treaties of the European Union and its Charter of Fundamental Rights), EU secondary law (such as the General Data Protection Regulation, the Product Liability Directive, the Regulation on the Free Flow of Non-Personal Data, anti-discrimination Directives, consumer law and Safety and Health at Work Directives), the UN Human Rights treaties and the Council of Europe conventions (such as the European Convention on Human Rights), and numerous EU Member State laws. Besides horizontally applicable rules, various domain-specific rules exist that apply to particular AI applications (such as for instance the Medical Device Regulation in the healthcare sector)” (Ibid., p. 6).

valores éticos⁶⁶⁰; e c) **robustez (confiabilidade)**, do ponto de vista técnico e em relação aos aspectos sociais.⁶⁶¹

A normativa europeia a respeito da Inteligência Artificial surgiu a partir de um contexto evolutivo das regras da Convenção do Conselho da Europa para a Proteção das Pessoas relativamente ao Tratamento Automatizado de Dados de Caráter Pessoal (Convenção 108 do Conselho da Europa).⁶⁶² Essa normativa, ao regulamentar os direitos fundamentais à vida privada e à proteção de dados pessoais (art. 7º e art. 8º da Carta de Direitos Fundamentais da União Europeia⁶⁶³), buscou garantir um maior nível de proteção às pessoas em relação ao tratamento automatizado de seus dados pessoais mediante “armazenamento de dados, realização de cálculos lógicos e/ou aritméticos operações sobre esses dados, sua alteração, apagamento, recuperação ou disseminação”⁶⁶⁴.

Já no *GDPR (General Data Protection Regulation)* da União Europeia, Regulamento (UE) 2016/679, o direito à explicabilidade pode ser inferido a partir do direito à obtenção de informações úteis para a compreensão da lógica computacional utilizada para a produção de decisões automatizadas (art. 14º, nº 2, alínea “g”⁶⁶⁵, e art. 15º, nº 1, alínea “h”), especialmente pelo fato de que em temas de “interesse

⁶⁶⁰ “Achieving Trustworthy AI requires not only compliance with the law, which is but one of its three components. Laws are not always up to speed with technological developments, can at times be out of step with ethical norms or may simply not be well suited to addressing certain issues. For AI systems to be trustworthy, they should hence also be ethical, ensuring alignment with ethical norms” (Ibid., p. 6-7).

⁶⁶¹ “Even if an ethical purpose is ensured, individuals and society must also be confident that AI systems will not cause any unintentional harm. Such systems should perform in a safe, secure and reliable manner, and safeguards should be foreseen to prevent any unintended adverse impacts. It is therefore important to ensure that AI systems are robust. This is needed both from a technical perspective (ensuring the system’s technical robustness as appropriate in a given context, such as the application domain or life cycle phase), and from a social perspective (in due consideration of the context and environment in which the system operates). Ethical and robust AI are hence closely intertwined and complement each other. The principles put forward in Chapter I, and the requirements derived from these principles in Chapter II, address both components” (Ibid., p. 7).

⁶⁶² UNIÃO EUROPEIA. **Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data**. Disponível em: <https://rm.coe.int/1680078b37>. Acesso em: 04 jan. 2023.

⁶⁶³ UNIÃO EUROPEIA. **Carta de Direitos Fundamentais da União Europeia**. Disponível em: https://www.europarl.europa.eu/charter/pdf/text_pt.pdf. Acesso em: 03 jan. 2023.

⁶⁶⁴ Tradução livre de: “c ‘automatic processing’ includes the following operations if carried out in whole or in part by automated means: storage of data, carrying out of logical and/or arithmetical operations on those data, their alteration, erasure, retrieval or dissemination” (UNIÃO EUROPEIA. **Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data**).

⁶⁶⁵ “g) A existência de decisões automatizadas, incluindo a definição de perfis referida no artigo 22º, n.ºs 1 e 4, e, pelo menos nesses casos, informações úteis relativas à lógica subjacente, bem como a importância e as consequências previstas de tal tratamento para o titular dos dados” (UNIÃO EUROPEIA. **Regulamento (UE) 2016/679 do Parlamento Europeu e do Conselho, de 27 de abril de 2016**. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:32016R0679&from=PT>. Acesso em: 05 jan. 2023).

público” (onde pode ser incluída a prestação jurisdicional) é inaplicável direito à oposição ao tratamento automatizado de dados pessoais (art. 21º, nº 6, do *GDPR*⁶⁶⁶).

Percebe-se, ademais, que desde que haja revisão humana sobre as decisões tomadas com o auxílio de sistemas de Inteligência Artificial (premissa para a inclusão de IA no auxílio às decisões judiciais, conforme relatado nesta obra), não há decisão “exclusivamente com tratamento automatizado” que possibilitaria cogitar a configuração do direito à oposição (art. 22º, nº 1, do *GDPR*⁶⁶⁷). Além disso, a Diretiva (UE) 2016/680 também embasou as Diretrizes Éticas para IA Confiável, uma vez que seu art. 11º limita a utilização de decisões totalmente automatizadas (sem revisão humana) para a definição de situações sensíveis da vida em sociedade.⁶⁶⁸

⁶⁶⁶ “1. O titular dos dados tem o direito de se opor a qualquer momento, por motivos relacionados com a sua situação particular, ao tratamento dos dados pessoais que lhe digam respeito com base no artigo 6.o, n.o 1, alínea e) ou f), ou no artigo 6.o, n.o 4, incluindo a definição de perfis com base nessas disposições. O responsável pelo tratamento cessa o tratamento dos dados pessoais, a não ser que apresente razões imperiosas e legítimas para esse tratamento que prevaleçam sobre os interesses, direitos e liberdades do titular dos dados, ou para efeitos de declaração, exercício ou defesa de um direito num processo judicial.

2. Quando os dados pessoais forem tratados para efeitos de comercialização direta, o titular dos dados tem o direito de se opor a qualquer momento ao tratamento dos dados pessoais que lhe digam respeito para os efeitos da referida comercialização, o que abrange a definição de perfis na medida em que esteja relacionada com a comercialização direta.

3. Caso o titular dos dados se oponha ao tratamento para efeitos de comercialização direta, os dados pessoais deixam de ser tratados para esse fim.

4. O mais tardar no momento da primeira comunicação ao titular dos dados, o direito a que se referem os n.ºs 1 e 2 é explicitamente levado à atenção do titular dos dados e é apresentado de modo claro e distinto de quaisquer outras informações.

5. No contexto da utilização dos serviços da sociedade da informação, e sem prejuízo da Diretiva 2002/58/CE, o titular dos dados pode exercer o seu direito de oposição por meios automatizados, utilizando especificações técnicas.

6. Quando os dados pessoais forem tratados para fins de investigação científica ou histórica ou para fins estatísticos, nos termos do artigo 89.º, nº 1, o titular dos dados tem o direito de se opor, por motivos relacionados com a sua situação particular, ao tratamento dos dados pessoais que lhe digam respeito, salvo se o tratamento for necessário para a prossecução de atribuições de interesse público” (Id.)

⁶⁶⁷ “1. O titular dos dados tem o direito de não ficar sujeito a nenhuma decisão tomada exclusivamente com base no tratamento automatizado, incluindo a definição de perfis, que produza efeitos na sua esfera jurídica ou que o afete significativamente de forma similar” (Id.)

⁶⁶⁸ “1. Os Estados-Membros preveem a proibição de decisões tomadas exclusivamente com base no tratamento automatizado, incluindo a definição de perfis, que produzam efeitos adversos na esfera jurídica do titular dos dados ou que o afetem de forma significativa, a menos que sejam autorizadas pelo direito da União ou do Estado-Membro ao qual o responsável pelo tratamento está sujeito, e que o direito da União ou desse Estado-Membro preveja garantias adequadas dos direitos e liberdades do titular dos dados, pelo menos o direito de obter a intervenção humana do responsável pelo tratamento.

2. As decisões a que se refere o nº 1 do presente artigo não se baseiam nas categorias especiais de dados pessoais a que se refere o artigo 8º, a não ser que sejam aplicadas medidas adequadas para salvaguardar os direitos e liberdades e os legítimos interesses do titular.

3. Em conformidade com o direito da União, são proibidas as definições de perfis que conduzam à discriminação de pessoas singulares com base nas categorias especiais de dados pessoais a que se refere o artigo 10º” (UNIÃO EUROPEIA. **Diretiva (UE) 2016/680 do Parlamento Europeu e do Conselho, de 27 de abril de 2016.** Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:32016L0680&from=HR>. Acesso em: 05 jan. 2023).

Nesse enfoque, na União Europeia há o estabelecimento do princípio da explicabilidade como forma de construir e manter a confiança dos usuários nos sistemas de IA mediante processos transparentes que explicitem as capacidades e os propósitos dos sistemas de IA, dentro do que é tecnicamente possível. Isso se reverte no dever dos sistemas de IA serem explicáveis às pessoas afetadas direta e indiretamente por suas decisões, a fim de possibilitar a contestação ou pedido de revisão.⁶⁶⁹ Todavia, a referida normativa editada pela Comissão Europeia prevê o seguinte:

“Uma explicação de por que um modelo gerou uma determinada saída ou decisão (e qual combinação de fatores de entrada contribuíram para isso) nem sempre é possível. Esses casos são chamados de algoritmos ‘caixa preta’ e requerem atenção especial. Nessas circunstâncias, outras medidas de explicabilidade (por exemplo, rastreabilidade, auditabilidade e comunicação transparente sobre as capacidades do sistema) podem ser necessárias, desde que o sistema como um todo respeite direitos fundamentais.”⁶⁷⁰

Ou seja, conforme já trabalhado nesta obra ao se delinear as características diferenciadoras da explicabilidade e da interpretabilidade (vide subcapítulo 2.6), a recomendação da União Europeia destaca que o “grau em que a explicabilidade é necessária depende muito do contexto e da gravidade das consequências se essa saída for errônea ou imprecisa”⁶⁷¹, o que também se relaciona com os princípios de transparência e *accountability*⁶⁷² (vide subcapítulos 2.3, 2.6 e 2.7).

⁶⁶⁹ UNIÃO EUROPEIA. **Ethics Guidelines for Trustworthy AI**, p. 13.

⁶⁷⁰ Tradução livres de: “An explanation as to why a model has generated a particular output or decision (and what combination of input factors contributed to that) is not always possible. These cases are referred to as ‘black box’ algorithms and require special attention. In those circumstances, other explicability measures (e.g. traceability, auditability and transparent communication on system capabilities) may be required, provided that the system as a whole respects fundamental rights” (Id.).

⁶⁷¹ Tradução livre de: “The degree to which explicability is needed is highly dependent on the context and the severity of the consequences if that output is erroneous or otherwise inaccurate” (Id.).

⁶⁷² “Deve haver transparência e divulgação responsável em torno de sistemas de IA para garantir que as pessoas entendam os resultados baseados em IA e possam questioná-los. Isso inclui *accountability* e registro de dados e documentação do *design*. O modelo de transparência e auditabilidade deve, sempre que possível, envolver um processo de certificação de boas práticas a ser ofertado pelos entes envolvidos (academia, indústria e governo)” (HARTMANN PEIXOTO, Fabiano. **Inteligência Artificial e Direito: Convergência Ética e Estratégica**, p. 156).

Veja-se, ainda, que, neste momento⁶⁷³, a União Europeia discute a “Proposta de Diretiva do Parlamento Europeu e do Conselho Relativa à Adaptação das Regras de Responsabilidade Civil Extracontratual à Inteligência Artificial”⁶⁷⁴, que busca estabelecer uma presunção de causalidade para os casos em que se configurarem danos emergentes de decisões tomadas com o auxílio de IA, porém o ônus de adequada explicabilidade⁶⁷⁵ seria das pessoas, empresas e entidades que se utilizem da referida tecnologia em seus processos decisórios. A referida proposta normativa, embora a princípio se destine aos casos de responsabilidade civil extracontratual, busca estabelecer uma regra de ônus da prova que, por analogia, poderá ser aplicada em qualquer caso em envolva o uso de Inteligência Artificial no trato de direitos fundamentais, desde que seja respeitado o estado da técnica da explicabilidade (vide subcapítulo 2.6).

Já nos G20⁶⁷⁶ *AI Principles*, verifica-se uma preocupação com a transparência e a explicabilidade, de maneira que o desenvolvimento de sistemas de IA se pautem na divulgação de informações significativas e apropriadas ao contexto de utilização e ao estado da arte da tecnologia, a fim de: a) promover uma **compreensão geral dos sistemas de IA**; b) possibilitar a **conscientização das pessoas interessadas** a respeito de suas interações com sistemas de IA; c) permitir que os afetados por um sistema de IA **tenham a possibilidade entender os resultados (outcomes)**; e d) possibilitar que os negativamente afetados por um sistema de IA possam **contestar**

⁶⁷³ Para fins de registro, este trecho da obra foi escrito em 27/12/2022. Assim, como o tema relativo à normatização da IA está em ebulição pelo mundo, o registro da data exata é necessário para contextualizar o leitor e possibilitar futuras atualizações por este autor.

⁶⁷⁴ UNIÃO EUROPEIA. **Proposta de Diretiva do Parlamento Europeu e do Conselho Relativa à Adaptação das Regras de Responsabilidade Civil Extracontratual à Inteligência Artificial**. [2022]. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:52022PC0496&from=EN>. Acesso em: 27 dez. 2022.

⁶⁷⁵ “(28) A presunção de causalidade pode igualmente aplicar-se a sistemas de IA que não sejam sistemas de IA de risco elevado, uma vez que o demandante pode encontrar dificuldades excessivas em termos de prova. Por exemplo, tais dificuldades podem ser apreciadas à luz das características de determinados sistemas de IA, como a autonomia e a opacidade, que, na prática, tornam muito difícil explicar o funcionamento interno do sistema de IA, afetando negativamente a capacidade do demandante em provar o nexo de causalidade entre o facto culposos do demandado e o resultado da IA. O tribunal nacional deve aplicar a presunção quando o demandante se encontra numa posição de dificuldade excessiva para provar o nexo de causalidade, uma vez que lhe é exigido que explique de que forma o sistema de IA foi levado, por ação ou omissão humana que constitui um facto culposos, a produzir o resultado ou a criar a incapacidade de produzir um resultado que deu origem ao dano. No entanto, não deve ser exigido ao demandante que explique as características do sistema de IA em causa nem a forma como estas características dificultam o estabelecimento do nexo de causalidade” (Ibid., p. 25).

⁶⁷⁶ O G20 é composto por Estados Unidos, Alemanha, França, Canadá, Itália, Japão, Reino Unido, Rússia, África do Sul, Arábia Saudita, Argentina, Austrália, Brasil, China, Coreia do Sul, Índia, Indonésia, México e Turquia.

os resultados com base em informações de fácil compreensão sobre os fatores e a lógica que serviram de base para a previsão, a recomendação ou a decisão.⁶⁷⁷

Por sua vez, a *G7's Charlevoix Common Vision for the Future of AI* define como linha estratégica para os países do referido grupo⁶⁷⁸ a promoção do investimento em pesquisa e desenvolvimento em IA de maneira a gerar confiança pública na tecnologia, haja vista a existência de questões críticas da Inteligência Artificial relacionadas com a *accountability*, a garantia, a responsabilidade, a segurança, a proteção e o risco de uso indevido.⁶⁷⁹

Também tendo em vista o mundo globalizado e as relações existentes entre o Brasil e os Estados Unidos da América, dentre os objetivos dos *AI Principles* do Departamento de Defesa dos EUA⁶⁸⁰ está o princípio da **rastreabilidade**, que estabelece que a engenharia de Inteligência Artificial voltada para a decisão orientada por dados (*DoD*) deve garantir que especialistas técnicos tenham uma **compreensão adequada da tecnologia, dos processos de desenvolvimento e dos métodos operacionais de seus sistemas de IA**. Essa compreensão passa por metodologias transparentes e auditáveis, registro das fontes de dados, estabelecimento procedimentos de *design* e documentação de todos os passos de desenvolvimento dos sistema de IA.⁶⁸¹

Ainda no contexto da realidade dos EUA, surgiu a proposta legislativa relacionada ao *"Blueprint for an AI Bill of Rights"*⁶⁸² editada pela Casa Branca no ano de 2022, que traz consigo a tentativa do estabelecimento de um amplo direito à explicabilidade, desde o desenvolvimento de sistemas que tenham a eles integrados a possibilidade de grau de explicabilidade até a pesquisa para a melhoria da interpretabilidade de modelos de IA⁶⁸³, por exemplo.

⁶⁷⁷ G20. Op. cit.

⁶⁷⁸ O G7 é composto por Estados Unidos, Alemanha, França, Canadá, Itália, Japão e Reino Unido.

⁶⁷⁹ G7. Op. cit.

⁶⁸⁰ ESTADOS UNIDOS DA AMÉRICA. **AI Principles: Recommendations on the Ethical Use of Artificial Intelligence by the Department of Defense**.

⁶⁸¹ Ibid., p. 8.

⁶⁸² ESTADOS UNIDOS DA AMÉRICA. **Blueprint for an AI Bill Of Rights**. [2022]. Disponível em: <https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf>. Acesso em: 27 dez. 2022.

⁶⁸³ *"The National Institute of Standards and Technology (NIST) is developing a risk management framework to better manage risks posed to individuals, organizations, and society by AI.19 The NIST AI Risk Management Framework, as mandated by Congress, is intended for voluntary use to help incorporate trustworthiness considerations into the design, development, use, and evaluation of AI products, services, and systems. The NIST framework is being developed through a consensusdriven, open, transparent, and collaborative process that includes workshops and other opportunities to provide*

Após a visão geral sobre o direito à explicabilidade em diplomas normativos e propostas de regulamentação supranacionais, transnacionais e estrangeiras, deve-se delimitar o tema em relação ao sistema de regional de proteção aos Direitos Humanos, no qual se observa a influência da Convenção Americana de Direitos Humanos (1969)⁶⁸⁴ sobre o tema do direito à explicabilidade, na medida em que o art. 3º da Convenção prevê que toda “pessoa tem direito ao reconhecimento de sua personalidade jurídica”.⁶⁸⁵ O referido direito encontra lastro no princípio de proteção à dignidade, que está previsto no art. 11, nº 1, da Convenção Americana: “Toda pessoa tem direito ao respeito de sua honra e ao reconhecimento de sua dignidade”.⁶⁸⁶

Nesse sentido, a normatividade dos direitos à dignidade e à liberdade individual (na forma prevista na Convenção Americana de Direitos Humanos, o que também passa pela interação com o art. 1º, III, e com o art. 5º, *caput*, da Constituição Federal) é intermediada pela observância do direito fundamental à **autodeterminação informativa**, que, no plano infraconstitucional, é explicitado no art. 2º, II, da Lei 13.709/2018.⁶⁸⁷

Desse modo, o direito fundamental à autodeterminação informativa passa pelo reconhecimento de que o processamento automatizado de dados ameaça a capacidade de decisão autônoma do indivíduo a respeito da oportunidade, da extensão e das consequências do fornecimento de seus dados pessoais para terceiros, pois há o risco de perfilamento⁶⁸⁸ da sua personalidade por sistemas de decisão automatizada, o que, muitas vezes, ocorre sem possibilidade de posterior

input. The NIST framework aims to foster the development of innovative approaches to address characteristics of trustworthiness including accuracy, explainability and interpretability, reliability, privacy, robustness, safety, security (resilience), and mitigation of unintended and/or harmful bias, as well as of harmful uses. The NIST framework will consider and encompass principles such as transparency, accountability, and fairness during pre-design, design and development, deployment, use, and testing and evaluation of AI technologies and systems. It is expected to be released in the winter of 2022-23” (Ibid., p. 21).

⁶⁸⁴ BRASIL. **Decreto 678, de 06 de novembro de 1992.**

⁶⁸⁵ Id.

⁶⁸⁶ Id.

⁶⁸⁷ BRASIL. **Lei 13.709, de 14 de agosto de 2018.**

⁶⁸⁸ “As técnicas de perfilamento têm como objetivo determinar o que é relevante dentro de um determinado contexto, por exemplo, quem pode estar interessado em um determinado produto. Além disto, estas técnicas auxiliam na representatividade estatística, ou seja, na determinação da qualidade de amostra constituída de modo a corresponder à população no seio de qual ela é escolhida. Ou seja, busca-se generalizar a partir de uma amostra de indivíduos e dos seus respectivos interesses” (BOFF, Salete Oro; FORTES, Vinícius Borges; FREITAS, Cinthia Obladen de Almendra. **Proteção de Dados e Privacidade: Do Direito às Novas Tecnologias na Sociedade da Informação.** Rio de Janeiro: Lumen Juris, 2018, p. 161)

controle sobre a utilização e a correção no uso dos dados.⁶⁸⁹ Por outro lado, o direito à autodeterminação informativa não é ilimitado, uma vez que as informações pessoais presentes nos dados tratados através de sistemas automatizados indicam parte da realidade social, o que denota o interesse da sociedade na utilização pública de parte dos dados (observadas a necessidade inerente à proporcionalidade).⁶⁹⁰ Assim, os riscos do processamento automatizado de dados residem “mais na finalidade do processamento e nas possibilidades de processamento do que no tipo dos dados tratados”.⁶⁹¹

Portanto, em um reforço necessário do direito à explicabilidade inerente ao duplo canal de correção na aplicação do direito, pode-se perceber que três elementos definem o direito à autodeterminação informativa: a) o **poder de decisão** dado ao indivíduo sobre a coleta e a utilização de seus dados pessoais é parte do âmbito de proteção da autodeterminação informativa⁶⁹²; b) a **ausência de teor fixo e definido** sobre a proteção (o reconhecimento de sua relação a dignidade e a personalidade humanas), que retira o direito de uma esfera puramente privada⁶⁹³ e o insere num suporte fático abstrato típico de uma estrutura de princípio de direito público⁶⁹⁴; e c) a **possibilidade de atribuição pessoal do dado tratado (identificação do indivíduo)**, oportunidade em que cada registro de dado pessoal deve ser protegido⁶⁹⁵ no âmbito das possibilidades tecnicamente possíveis.

Desse modo, reconhece-se a necessidade do estabelecimento da eficácia horizontal (em face de entidades privadas que realizam o tratamento de dados pessoais)⁶⁹⁶ e da eficácia vertical (em face de entes estatais) do direito fundamental à

⁶⁸⁹ MENDES, Laura Schertel Ferreira. In: DONEDA, Danilo; SARLET, Ingo Wolfgang; MENDES, Laura Schertel Ferreira. **Estudos Sobre Proteção de Dados Pessoais**: Dados num Mundo em Transformação. São Paulo: Expressa, 2022, p. 36. E-book Kindle.

⁶⁹⁰ Id.

⁶⁹¹ Ibid., p. 38.

⁶⁹² Ibid., p. 39-40.

⁶⁹³ Ibid., p. 40.

⁶⁹⁴ ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 305.

⁶⁹⁵ MENDES, Laura Schertel Ferreira. Op. cit., p. 40.

⁶⁹⁶ “O direito à autodeterminação informativa, que garante o poder do indivíduo referente ao fornecimento e utilização dos seus dados pessoais, também age como norma de proteção, ou seja, como direito objetivo no âmbito privado e, nessa medida, deve ser levado em consideração pelos juízes em um caso concreto” (Ibid., p. 43-44).

autodeterminação informativa⁶⁹⁷, o que ocorre no contexto de um direito geral de personalidade⁶⁹⁸ com fundo Constitucional e de Direitos Humanos.

Nesse prisma, ainda que haja uma concepção inicialmente procedimental do direito à autodeterminação informativa, existem elementos substanciais inerentes aos riscos concretos à liberdade individual e ao processamento de informações, o que transita pela tensão entre o poder de decisão⁶⁹⁹ e a necessidade de tratamento de dados pessoais nas questões interesse coletivo (como no caso da prestação jurisdicional).

Além disso, no plano jurisprudencial, destaca-se o julgamento da Medida Cautelar na Ação Direta de Inconstitucionalidade 6387 (ADI 6387 MC), oportunidade em que o Supremo Tribunal Federal reconheceu a autodeterminação informativa como um direito fundamental inerente ao sistema constitucional brasileiro.⁷⁰⁰

⁶⁹⁷ “Por conseguinte, da dimensão objetiva do direito à autodeterminação informativa decorre não apenas um dever de proteção contra o conhecimento não autorizado por terceiros, mas também um dever de proteção contra o consentimento meramente aparente (ou fictício) quanto ao tratamento de dados. Dos deveres de proteção decorrentes do direito à autodeterminação informativa resulta, assim, um controle judicial específico também de contratos que regulam o tratamento de dados” (Ibid., p. 46-47).

⁶⁹⁸ “Logo, da mesma forma que no direito geral de personalidade, a relevância da autodeterminação informativa reside na flexibilidade oportunizada por uma concepção abstrata de proteção, pois se reporta ao poder de decisão do indivíduo, oferecendo grande” (Ibid., p. 50).

⁶⁹⁹ Ibid., p. 51.

⁷⁰⁰ “1. Decorrências dos direitos da personalidade, o respeito à privacidade e à autodeterminação informativa foram positivados, no art. 2º, I e II, da Lei nº 13.709/2018 (Lei Geral de Proteção de Dados Pessoais), como fundamentos específicos da disciplina da proteção de dados pessoais. 2. Na medida em que relacionados à identificação – efetiva ou potencial – de pessoa natural, o tratamento e a manipulação de dados pessoais hão de observar os limites delineados pelo âmbito de proteção das cláusulas constitucionais assecuratórias da liberdade individual (art. 5º, caput), da privacidade e do livre desenvolvimento da personalidade (art. 5º, X e XII), sob pena de lesão a esses direitos. O compartilhamento, com ente público, de dados pessoais custodiados por concessionária de serviço público há de assegurar mecanismos de proteção e segurança desses dados”. (...) “5. Ao não definir apropriadamente como e para que serão utilizados os dados coletados, a MP nº 954/2020 desatende a garantia do devido processo legal (art. 5º, LIV, da CF), na dimensão substantiva, por não oferecer condições de avaliação quanto à sua adequação e necessidade, assim entendidas como a compatibilidade do tratamento com as finalidades informadas e sua limitação ao mínimo necessário para alcançar suas finalidades. 6. Ao não apresentar mecanismo técnico ou administrativo apto a proteger, de acessos não autorizados, vazamentos acidentais ou utilização indevida, seja na transmissão, seja no tratamento, o sigilo, a higidez e, quando o caso, o anonimato dos dados pessoais compartilhados, a MP nº 954/2020 descumpra as exigências que exsurgem do texto constitucional no tocante à efetiva proteção dos direitos fundamentais dos brasileiros” (BRASIL. Supremo Tribunal Federal (Plenário). **Medida Cautelar em Ação Direta de Inconstitucionalidade 6387**. Defere medida cautelar para afastar os efeitos da Medida Provisória nº 954/2020 em razão da ofensa ao direito fundamental à autodeterminação informativa. Brasília, DF: Supremo Tribunal Federal, [2020]. Relator: Min. Gilmar Mendes, 07 de maio de 2020. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=754357629>. Acesso em: 28 dez. 2022).

A contextualização a respeito dos aspectos dogmáticos já passíveis de análise para o delineamento do direito à explicabilidade, inerente ao duplo canal de correção, torna possível a compreensão do alcance normativo do princípio da **transparência** da Inteligência Artificial que vem sendo discutido no Projeto de Lei nº 21/2020⁷⁰¹ (atualmente em trâmite no Senado Federal), o que se somará ao já referido art. 20 da Lei 13.709/2018 (LGPD).

Outro aspecto a ser considerado, frente à sua relação com alcance do direito à explicabilidade, é a necessária anonimização dos dados pessoais na fase de preparação dos dados (o que fica mais evidente no método CRISP-DM⁷⁰²), de maneira que as variáveis preditoras (aquelas que serão exploradas pelos algoritmos de *machine learning*) não contenham a possibilidade de identificação dos titulares dos dados. Destarte, tendo em vista que o tratamento dos dados para o treinamento de sistemas de IA judicial ocorre no contexto de políticas públicas do Poder Judiciário e de pesquisa (art. 7º, III e IV, art. 11, II, “b” e “c”, da Lei 13.709/2018), do ponto de vista dogmático, inexistente necessidade de prévia autorização do titular, uma vez que os dados anonimizados dispensam a colheita de autorização de qualquer pessoa.

Nesse sentido, realizado o processo de anonimização para transformação de dados pessoais em dados brutos não identificáveis, para fins legais o *dataset* deixa de conter “dados pessoais”, uma vez que os “dados anonimizados não serão considerados dados pessoais” (art. 12, *caput*, da Lei 13.709/2018), salvo em caso de reversão.

Assim, o direito à anonimização de dados pessoais (art. 18, IV, da Lei 13.709/2018) atua como uma garantia ao titular dos dados, a fim de que processo exploratório de dados, tipicamente necessário para a criação de modelos computacionais baseados em aprendizado de máquina, e o direito à explicabilidade (quando demande acesso à base de dados utilizada pela IA) não violem o direito fundamental à autodeterminação informacional.

Dessa forma, a anonimização prévia dos dados, antes de qualquer compartilhamento ou utilização para o desenvolvimento de sistema de Inteligência Artificial, previne a necessidade de aplicação da regra do §6º do 18 da Lei 13.709/2018, que define que o responsável deve informar, de maneira imediata, aos

⁷⁰¹ BRASIL. Senado Federal. **Projeto de Lei nº 21, de 2020.**

⁷⁰² Vide o Capítulo 5.

agentes de tratamento com os quais tenha realizado uso compartilhado de dados, “a anonimização ou o bloqueio dos dados, para que repitam idêntico procedimento”.⁷⁰³

Em suma, frente à anonimização dos dados pessoais a serem utilizados no desenvolvimento diversos riscos são prevenidos. Dessa forma, a fim de se garantir que a anonimização seja efetiva⁷⁰⁴, os dados armazenados já devem ter a identificação dos titulares descartada previamente (o que se relaciona aos dados desnecessários para o treinamento da IA). O único lugar onde deve ser possível identificar os titulares é o repositório original do banco de dados (por exemplo, os dados de identificação de pacientes devem estar disponíveis para os sistemas de saúde, não para a equipe de desenvolvimento de sistemas de IA judicial).

O direito à explicabilidade não garante o acesso à identificação pessoal dos titulares dos dados utilizados para o treinamento dos sistemas de Inteligência Artificial judicial, uma vez que, no conforme visto, isso causaria lesão ao direito à autodeterminação informativa, na perspectiva de proteção à dignidade humana e à privacidade dos titulares dos dados. Trata-se de um equilíbrio necessário para que os direitos coexistam no sistema de direitos fundamentais.

Logo, tendo em vista que o suporte fático abstrato do direito fundamental à explicabilidade encontra lastro na própria ideia de autodeterminação informativa e no princípio da transparência, para que exista explicabilidade das decisões judiciais auxiliadas por sistemas de Inteligência Artificial, deve existir:

a) a necessária oportunização de conhecimento ao indivíduo que tenha seus dados pessoais tratados no processo de modelagem de Inteligência Artificial voltada a sistema computacional de auxílio à decisão judicial;

b) o fornecimento de explicação de maneira compreensível à pessoa (explicação não técnica) a respeito da lógica, dos passos e do processo de

⁷⁰³ BRASIL. **Lei 13.709, de 14 de agosto de 2018.**

⁷⁰⁴ “Quanto às técnicas de anonimização, deve-se apontar para a estruturação clara de um pacote/programa de governança digital que vise efetivamente reduzir e minimizar os riscos e os impactos de incidentes de segurança, em especial no que concerne ao compartilhamento e ao cruzamento de dados.

Assim, os algoritmos devem servir, em regra, para a emancipação e, assim, para aumentar, complementar e capacitar as habilidades cognitivas, sociais, éticas e culturais dos seres humanos. Nesse ponto, oportunamente, entra em cena o debate sobre os limites da governança digital relacionada à responsabilidade algorítmica, que, em rigor, não será tratado direta e nem profundamente nesse artigo apesar de sua extrema significância” (SARLET, Gabrielle Bezerra Sales; RUARO, Regina Linden. A Proteção de Dados Sensíveis no Sistema Normativo Brasileiro sob o Enfoque da Lei Geral de Proteção de Dados (LGPD) – L.13.709/2018. **Revista Direitos Fundamentais & Democracia**, Curitiba, v. 26, n. 2, p. 81-106, mai./ago. 2021. DOI: 10.25192/issn.1982-0496.rdfd.v26i22172).

desenvolvimento da Inteligência Artificial utilizada para auxiliar a decisão judicial, dos aspectos do funcionamento do modelo de IA, desde que respeitado o limite da técnica (uma vez que a explicabilidade se dá em graus diversos a depender da tecnologia adotada, mas sempre no máximo alcance possível);

c) a possibilidade de desafio técnico ao sistema de Inteligência Artificial, mediante o fornecimento de informações disponíveis que possibilitem a interpretabilidade do sistema computacional por qualquer profissional técnico interessado (advindo da sociedade civil, da academia, de órgãos de controle ou contratado por qualquer parte interessada), sendo que:

1. em sistemas baseados em representação de conhecimento o acesso aos algoritmos deve ser integralmente garantido;

2. em sistemas baseados em aprendizado de máquina (*machine learning*) a base de dados anonimizada e os algoritmos devem ser disponibilizados;

3. em sistemas baseados em aprendizado profundo (*deep learning*) as bases de dados devem ser referidas em relação à integralidade de suas origens (possibilidade de rastreabilidade), ainda que haja baixa possibilidade de interpretabilidade;

d) a explicação de quem são os integrantes da equipe responsável pelo desenvolvimento, pela implantação e pela manutenção do sistema de IA judicial.

Portanto, a análise dos aspectos dogmáticos do direito à explicabilidade contribui para o desenvolvimento de um firme entrelace entre o Direito e a Ciência da Computação, de maneira que exista um mútuo aprimoramento científico que resulta em melhoria para a vida das pessoas, segundo os elementos trabalhados neste Capítulo. Além disso, essa cooperação tem suas bases em outros ingredientes, que são trabalhados nos próximos Capítulos.

3 UMA COOPERAÇÃO POSSÍVEL ENTRE INTELIGÊNCIA ARTIFICIAL E DECISÃO JUDICIAL

3.1 Aspectos da Objetividade dos Sistemas da IA Judicial

O desenvolvimento de sistemas de Inteligência Artificial voltados para o auxílio à decisão judicial possibilita uma maior objetividade e mais previsibilidade na aplicação do direito, o que se dá através do duplo canal de correção advindo da relação juiz-IA, descrito nesta obra.

Nesse caminho, a cooperação entre Inteligência Artificial e decisão judicial se aproveita das características que funcionam como catalizadoras dessa interação, uma delas é a objetividade advinda dos sistemas jurídicos e de suas normas, o que se passa a tratar em razão de sua influência no funcionamento do duplo canal de correção, em especial no que se refere à pretensão de correção normativa.

Uma norma jurídica é produto da escolha do legislador ao tentar alcançar as condutas humanas por meio de uma máquina preditiva construída no ordenamento jurídico⁷⁰⁵. Desse modo, um ordenamento jurídico (integrado pela normatividade administrativa, penal, civil, processual, etc.) é composto por diversos sistemas jurídicos⁷⁰⁶ que são recortes, com algum grau de coesão ou microssistemas dentro daquele contexto maior (o sistema de responsabilidade civil por danos ao consumidor ou a legislação de trânsito, por exemplo). Nesse sentido, existe um grau de objetividade intrínseco aos sistemas normativos⁷⁰⁷ que nasce do grau de consistência, do potencial de completude e da sua relativa independência.

⁷⁰⁵ Este autor adota a perspectiva de diferenciação conceitual entre ordenamento jurídico (que considera toda a legislação de um estado, por exemplo, e que guarda laços de unidade e sistematicidade mais fracos) e sistema jurídico (que considera a legislação aplicada a determinado setor do estado ou da sociedade, ou ainda a determinado ramo do Direito, dentro de uma compreensão de forte sistematicidade e coesão de suas normas).

⁷⁰⁶ *“En la teoría jurídica tradicional, y en las obras de filosofía jurídica, por ‘sistema jurídico’ o bien por ‘orden jurídico’ se entiende el derecho en su globalidad, la totalidad de las normas jurídicas. Pero, en la práctica, los jueces, abogados, juristas, etc., nunca están interesados en la totalidad del orden jurídico. Ellos siempre se ocupan de una selección (usualmente más bien pequeña) de disposiciones o normas jurídicas, que son relevantes para el problema que les interesa. En ‘Normative Systems’, por sistema jurídico se entiende un microssistema, no la totalidad del orden jurídico”* (BULYGIN, Eugenio. **Lógica Deóntica, Normas y Propositiones Normativas**, p. 53).

⁷⁰⁷ *“It can thus reasonably be expected that an approach to the traditional problems of legal philosophy in the light of modern logical and methodological investigations would produce interesting results. The notion of normative system seems particularly suitable for such purpose. On the one hand, although the structure of deductive systems with their formal properties (consistency, completeness and independence) is one of the main topics of modern foundational studies both in their purely formal*

Destaca-se que embora Bulygin e Alchourrón⁷⁰⁸ reconheçam a consistência, a completude e a independência como propriedades de um sistema jurídico, deve-se considerar que a consistência possui uma gradação, que a completude é uma pretensão nem sempre alcançável e que a independência é apenas relativa, na medida em que há a inter-relação entre os diversos sistemas jurídicos que fazem parte do ordenamento jurídico.⁷⁰⁹ Nesse sentido, Bulygin e Alchourrón reconhecem que existem sistemas que são inconsistentes, incompletos e/ou dependentes de outros sistemas⁷¹⁰, características que, por si só, não impedem o uso de lógica para emprestar o grau de objetividade dos sistemas jurídicos através de sistemas de Inteligência Artificial voltados ao auxílio da decisão judicial.

Em suma, um ordenamento jurídico é composto de diversos sistemas jurídicos, assim como um sistema jurídico é composto de diversas normas. Todavia, reconhecendo a separação entre texto e norma⁷¹¹, deve ser observado que cada norma tem certos atributos que a constituem. Bulygin, por exemplo, reconhece os **modais deônticos de proibição, obrigação e permissão** como os mais básicos atributos que uma norma pode conter, atributos que, na segunda fase de pensamento do autor, são tipicamente jurídicos ao se realizar a passagem da lógica deôntica para a lógica de normas, e desta para a **lógica de proposições normativas**.⁷¹²

Com foco na explicabilidade, inerente ao duplo canal de correção, os atributos de cada norma podem ser trabalhados na programação dos sistemas de Inteligência Artificial judicial de diferentes maneiras, em consideração ontológica ou pressuposição heurística a respeito dos fatos sob julgamento (mediante novos atributos de programação jurídica) ou mediante sistemas de aprendizagem de máquina (*machine learning*) especialmente desenvolvidos para tratar os dados fáticos. Isso possibilitará a interação fato-norma nos sistemas de IA judicial ou a partir deles, o que é inerente à aplicação do Direito, tema que será explorado nos Capítulos 4 e 5.

aspects (uninterpreted, syntactical systems), specifically normative systems have been studied only by deontic logicians as purely formal (logical) systems" (ALCHOURRÓN, Carlos E.; BULYGIN, Eugenio. Op. cit., p. 3).

⁷⁰⁸ Ibid., p. 22-27.

⁷⁰⁹ Por exemplo, a interpretação extensiva e a analogia são exemplos desta característica.

⁷¹⁰ ALCHOURRÓN, Carlos E.; BULYGIN, Eugenio. Op. cit., p. 128-130.

⁷¹¹ ÁVILA, Humberto. **Teoria dos Princípios**: da definição à aplicação dos princípios jurídicos. 5ª ed. São Paulo: Malheiros, 2006, p. 30-31.

⁷¹² BULYGIN, Eugenio. **Lógica Deôntica, Normas y Proposiciones Normativas**, p. 40-42.

Do ponto de vista de cada norma isoladamente considerada e dos atributos que cada norma carrega, em relação ao prévio carregamento de sentido, existem normatividades mais densas e outras de densidade rarefeita, e, entre tais extremos, há uma grande possibilidade de distintas finalidades das normas. Para uma abordagem a respeito da possibilidade de utilização de Inteligência Artificial para aplicação do direito, a compreensão desta característica é necessária. Essa característica pode ser denominada como **densidade semântica**. Nesse sentido, embora haja a possibilidade de textura aberta em determinados casos⁷¹³, o tipo de normas com maior densidade semântica são as regras⁷¹⁴, que têm um maior número de atributos voltados à definição de sentido e à aplicação do direito⁷¹⁵, mesmo para os casos onde a subsunção se dá pela análise da probabilidade em relação à generalização.⁷¹⁶ Já, no plano normativo (ainda sem a consideração do suporte fático⁷¹⁷), os princípios têm poucos atributos voltados à imediata definição de sentido.⁷¹⁸ Ou seja, embora existam posicionamentos que rejeitem a objetividade dos sistemas jurídicos⁷¹⁹ como uma **característica que existe num gradiente**, observa-se que o Direito assim se desenha.

⁷¹³ “Open texture is the ineliminable possibility of vagueness, the ineradicable contingency that even the most seemingly precise term might, when it confronts an instance unanticipated when the term was defined, become vague with respect to that instance” (SCHAUER, Frederick. **Playing by the Rules: A Philosophical Examination of Rule-Based Decision-Making in Law and Life**. Oxford: Oxford University Press, 1991, p. 36. E-book Kindle).

⁷¹⁴ “As with description, therefore, prescriptive generalization commonly starts with a particular, takes it as an exemplar of some more general category, and then searches for the property of the particular that is causally relevant to the occurrence of the more general category” (Ibid., p. 25-26).

⁷¹⁵ “The justification thus determines which among logically equivalent generalizations from some particular precipitating event will be selected as the factual predicate of the ensuing rule” (Ibid., p. 26).

⁷¹⁶ “The factual predicate thus represents a set of facts whose existence indicates an increased likelihood of the occurrence of the justification, and whose (effective) prohibition will consequently decrease the likelihood (or incidence) of the evil against which the rule is aimed. Similarly, if the rule is seen to have been established in pursuit of a positive goal rather than diminution of an evil, the existence of the condition represented by the factual predicate will, when the probabilistic relationship exists, increase the likelihood of the positive goal toward which the rule is targeted. The key in either case is the probabilistic relationship, explaining why ‘No dogs allowed’ is an appropriate response to the Anguses of the future, while at the same time recognizing both that not all dogs will cause annoying disruptions, and that not all annoying disruptions will be caused by dogs” (Ibid., p. 30-31).

⁷¹⁷ Vide o Capítulo 4.

⁷¹⁸ Do ponto de vista linguístico, os princípios têm um espaço maior de adaptabilidade para a atribuição de novos sentidos e novas possibilidades de aplicação, maleabilidade que é sua característica. Por outro lado, respeitadas as características típicas dessa espécie normativa, a computação jurídica de princípios é possível através dos atributos e elementos que os definem. Para aprofundamento, o tema será abordado nos subcapítulos 4.1 e 4.3.1 da presente obra.

⁷¹⁹ “A maioria quase absoluta dos enunciados jurídicos, muito mais ainda a totalidade das normas constitucionais de direitos fundamentais, não satisfazem nem de longe as exigências que a lógica das inferências deve formular com relação a premissas maiores” (MÜLLER, Friedrich. **Teoria Estruturante do Direito**. 2ª ed. rev., atual. e amp. Tradução: Peter Naumann e Eurides Avance de Souza. São Paulo: Revista dos Tribunais, 2009, p. 22).

Dessa forma, em comparação com os princípios, a aplicação de uma regra pode mais facilmente ser inserida num algoritmo de lógica computacional, em programação procedural ou para garantir um fiel ponto de referência para a eficiência de modelos de IA (acurácia, sensibilidade e precisão⁷²⁰) treinados em grandes *datasets*. Uma regra diz, com alto grau de exatidão e previsibilidade, qual é o objeto ideal para a implementação de IA na sua aplicação. Tratar do tema exige um grau de complexidade elevado, mediante as aproximações possíveis entre Direito e Ciência da Computação. Dessa maneira, o desenvolvimento de sistemas de IA judicial deve levar em conta as pertinentes críticas à tentativa de congelamento do direito⁷²¹, para que tal equívoco seja prevenido desde a fase de projeto, pois num contexto de lógicas não-monotônicas⁷²² se verifica a possibilidade de derrotabilidade das normas jurídicas.⁷²³

Nesse momento, o que pode ser destacado é que as normas jurídicas apresentam grau de objetividade que determina a ideia de previsibilidade. Assim, a previsibilidade da futura aplicação do direito tem muito de sua força no estabelecimento de regras jurídicas, pois, como lembra Hans Kelsen, com as regras jurídicas surge “dever de conduzir-se de certa maneira”⁷²⁴ (dever jurídico), levando-se em conta que esse dever é sinônimo da obrigação prescrita objetivamente na norma jurídica.⁷²⁵

Nesse sentido, Kelsen destaca que, ainda que seja reconhecida a função criadora do Direito (em maior grau no contexto da *Common Law* e em menor grau no

⁷²⁰ Os conceitos de acurácia, sensibilidade e precisão estão descritos no subcapítulo 5.6.

⁷²¹ A “compreensão falha da norma jurídica como juízo hipotético transformado em proposição de lógica formal, como ordem a ser igualada ao seu texto linguístico, como premissa maior a ser tratada segundo regras silogísticas, propõe, a partir da concretização do direito em formulação mais aguda, a questão de uma definição mais adequada da relação entre direito e realidade e com isso simultaneamente a questão de uma compreensão sustentável da norma, bem como da ciência jurídica como ciência das normas” (MÜLLER, Friedrich. Op. cit., p. 22).

⁷²² CELLA, José Renato Gaziero. Op cit., p. 156-158.

⁷²³ “Sem embargo da falta de univocidade do termo ‘derrotabilidade’, é possível afirmar a existência de um núcleo comum atribuído ao seu conceito, consistente na ideia segundo a qual a consequência da norma jurídica pode ser derrotada, afastada, não-aplicada, em razão da existência de um fato, interpretação ou circunstância com ela incompatível (Cf., capítulo 5, infra), pode-se falar em fatos impeditivos, modificativos ou extintivos capazes de infirmar a norma ‘*prima facie*’, derrotando-a a fim de permitir a aplicação de uma norma excepcional, diferente daquela prevista *a priori* a partir da literalidade textual” (VASCONCELLOS, Fernando Andreoni. **Hermenêutica Jurídica e Derrotabilidade**. Curitiba: Juruá, 2010, p. 57).

⁷²⁴ KELSEN, Hans. **Teoria Pura do Direito**. 8.^a ed, p. 138.

⁷²⁵ Id.

sistema da *Civil Law*), o julgador não é o agente criador primário das normas, pois é um coadjuvante nessa missão, sob pena de total renúncia à segurança jurídica.⁷²⁶

Dessa forma, a Inteligência Artificial judicial, desde que pautada na busca da objetividade da norma jurídica, tem o potencial de fomentar o funcionamento do duplo canal de correção ao exigir a justificação racional pelo julgador para haja a superação de decisões propostas pelo direito algorítmico, pois, em respeito ao dever judicial de sinceridade⁷²⁷, os argumentos algorítmicos não podem ser ocultados por um juiz receoso de enfrentá-los. Isso ocorre porque mesmo em um contexto pós-positivista, o respeito à normatividade moldada pelo poder constituinte e pelo legislador é parte do Direito⁷²⁸, o que não impede a pretensão de correção, mas, ao contrário, a exige ainda mais e a refina.

Todavia, quanto menor a densidade normativa (princípios⁷²⁹ e textura aberta da linguagem⁷³⁰), menor será a objetividade na aplicação e mais cuidado deve se ter na utilização da Inteligência Artificial como um auxiliar na aplicação do Direito.

A respeito das vantagens da objetividade⁷³¹, por exemplo, existem milhares de ações previdenciárias que esperam julgamento, com casos que se repetem, cuja utilização de Inteligência Artificial pode otimizar o recebimento do crédito pelas pessoas que têm direito. O direito previdenciário é um microsistema regido majoritariamente por regras de aplicação por subsunção, sendo típico objeto para projetos de Inteligência Artificial judicial.

Por outro lado, a aplicação de Inteligência Artificial em casos de indenização por dano moral, por exemplo, enfrenta um mundo fático (ontológico) de alcance mais

⁷²⁶ “É o sistema da livre descoberta do Direito, sistema que já Platão propôs para o seu Estado ideal. Em consequência da radical descentralização da criação do Direito que lhe vai ligada, este sistema caracteriza-se pela sua grande flexibilidade, mas, em contrapartida, renuncia totalmente à segurança jurídica. Com efeito, sob uma tal ordem jurídica, os indivíduos não podem prever as decisões dos casos concretos em que participam como demandantes e demandados, acusadores e acusados. Por isso, não podem de forma nenhuma saber de antemão o que é que lhes é juridicamente proibido ou permitido, para que é que têm ou não têm poder ou competência. Somente podem conhecer isso através da decisão em que lhes é aplicada uma pena ou em que são absolvidos da acusação, em que a ação é rejeitada ou recebe provimento” (Ibid., p. 280).

⁷²⁷ “(1.2) Todo falante só pode afirmar aquilo em que ele mesmo acredita” (ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 168).

⁷²⁸ ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 106 e p. 308.

⁷²⁹ Id., p. 90.

⁷³⁰ HART, Herbert A. L. Op. cit., p. 164.

⁷³¹ PEREIRA, João Sérgio dos Santos Soares. **A Padronização Decisória na Era da Inteligência Artificial**: uma possível leitura hermenêutica e da autonomia do direito. Belo Horizonte: Casa do Direito, 2021, p. 347.

desafiador para o direito algorítmico, pelo fato de que o domínio⁷³² (responsabilidade civil) é amplíssimo, pois existe um grau de discricionariedade do julgador (maior ou menor, a depender do caso) na fixação das indenizações mediante a aplicação de critérios razoavelmente abertos (previstos em lei e nos precedentes). Nesse exemplo da responsabilidade civil o que se pode esperar de um sistema de Inteligência Artificial judicial, em razão dos limites da tecnologia, da ética e segurança necessárias no uso de IA, é o objetivo balizamento da sugestão de decisão nos limites das regras aplicáveis⁷³³ e o auxílio na composição do valor do *quantum* indenizatório mediante a sugestão de opções de valores de indenização segundo critérios advindos de precedentes judiciais e do histórico de decisões do próprio órgão julgador.

Para os mais céticos, ficam os exemplos do *GitHub Copilot*⁷³⁴ e do *Alpha Code*⁷³⁵, aplicações de Inteligência Artificial que atuam na sugestão de algoritmos para desenvolvedores⁷³⁶, os quais têm a decisão final sobre se é adequada a utilização de uma das sugestões, bem como qual das sugestões de algoritmos é a melhor para o caso. Uma abordagem análoga pode ser utilizada nos sistemas de IA judicial, sem necessária perda da oportunidade de decisão do julgador.

Portanto, a objetividade da IA judicial não surge num passe mágica, pois para sua implementação deve ser empreendido esforço justificatório maior do que normalmente os juízes fazem, uma vez que os atributos da norma jurídica desenhados no direito algorítmico têm limitações⁷³⁷, cabendo ao órgão julgador exercer seu papel constitucional.⁷³⁸ Mais uma vez, não custa lembrar a necessária interação humano-máquina como um elemento positivo e agregador na pretensão de correção.

⁷³² O conceito de domínio aqui se refere à significado atribuído pela ciência da computação.

⁷³³ “Primeiro: nunca subestime o poder da inércia. Segundo: é possível utilizar esse poder em benefício próprio” (THALER, Richard H.; SUNSTEIN, Cass R. Op. cit., posição 185).

⁷³⁴ GITHUB. **GitHub Copilot: Your AI pair programmer**. [2021]. Disponível em: <https://copilot.github.com>. Acesso em: 04 fev. 2022.

⁷³⁵ DEEPMIND. **AlphaCode**. [2022]. Disponível em: <https://alphacode.deepmind.com>. Acesso em: 04 fev. 2022.

⁷³⁶ LI, Yujia et al. **Competition-Level Code Generation with AlphaCode**. Disponível em: https://storage.googleapis.com/deepmind-media/AlphaCode/competition_level_code_generation_with_alphacode.pdf. Acesso em: 04 fev. 2022.

⁷³⁷ “A premissa de um dos erros mais fundamentais do positivismo na ciência jurídica, a compreensão e o tratamento da norma jurídica como algo que repousa em si e preexistente, é a separação da norma e dos fatos, do direito e da realidade” (MÜLLER, Friedrich. **Teoria Estruturante do Direito**, p. 16).

⁷³⁸ “A estrutura da norma e a normatividade apresentar-se-ão de formas tão distintas, de acordo com as diferenças materiais entre as disciplinas jurídicas, bem como de acordo com as diferenças materiais entre os diferentes tipos de normas. Por isso o direito público e direito constitucional são em grau especificamente elevado direito ‘político’; eles se referem à constituição da sociedade politicamente organizada e do ordenamento jurídico considerado na sua totalidade e as suas prescrições não se respaldam em normas de direito positivo de grau hierárquico superior” (Ibid., p. 45).

Esses exemplos demonstram que uma abordagem da Inteligência Artificial na busca de otimização da objetividade e da saturação argumentativa⁷³⁹ deve ocorrer sem que haja a ilusão de total neutralidade.⁷⁴⁰

Dessa maneira, a objetividade dos sistemas jurídicos deve ser utilizada como forma de desenvolver os sistemas de Inteligência Artificial judicial com foco na previsibilidade e na pretensão de correção, de maneira que o Direito fomente legitimidade à Inteligência Artificial judicial.

3.2 Robô Processual, Regra Padrão e Esquecimento: Heurísticas nos Atos Processuais de Maior Simplicidade

É sempre desafiador falar sobre processo, para além dos processualistas (civis, penais, trabalhistas, etc.) tal tema desperta interesse de juristas de variados ramos do direito, sendo que o direito processual é entrelaçado ao direito constitucional e à sua racionalidade.

Antes de se falar em Inteligência Artificial processual, é natural que a abordagem sobre o tema tenha um passo inicial na análise da regra do inciso XIV do art. 93 da Constituição Federal (com redação dada pela Emenda Constitucional 45/2004), a qual prevê que “os servidores [do Poder Judiciário] receberão delegação para a prática de atos de administração e atos de mero expediente sem caráter decisório”.

O Supremo Tribunal Federal (mediante voto do Ministro Dias Toffoli no RE 820433 AgR) ressaltou a importância da regra do 93, XIV, da Constituição Federal ao apontar que há uma “prerrogativa constitucional de os servidores receberem delegação para a prática de atos de administração e de mero expediente sem caráter decisório”.⁷⁴¹

⁷³⁹ “Um regresso ao infinito apenas se poderá evitar se a fundamentação se interrompe em algum momento e se substitui por uma decisão que já não se tem de fundamentar” (ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 160).

⁷⁴⁰ “Faz-se necessário observar que a neutralidade judicial não apenas é enganosa, como também constitui uma fonte (re)produtora de injustiças. Isso porque a neutralidade desconsidera o fato de que os direitos que existem são, eles próprios, produtos de leis” (LORENZETTO, Bruno Meneses. **Os Caminhos do Constitucionalismo Para a Democracia**. Tese (Doutorado em Direito). Universidade Federal do Paraná (UFPR), Curitiba, 2014, p. 275).

⁷⁴¹ BRASIL. Supremo Tribunal Federal (Plenário). **Agravo Regimental no Recurso Extraordinário 820.433**. Constitucionalidade de desersão de recurso reconhecida pelo Superior Tribunal de Justiça mediante a aplicação do art. 93, XIV, da Constituição Federal. Brasília, DF: Supremo Tribunal Federal, [2016]. Relator: Min. Dias Toffoli, 17 de março de 2016. Disponível em:

Já no âmbito do Superior Tribunal de Justiça há precedente (REsp 1177612/SP), no sentido de que os atos meramente ordinatórios podem ser delegados porque “não possuem a potencialidade de causar prejuízo a qualquer das partes envolvidas nos processos”, bem como que tais atos “facilitam a celeridade do processo e podem ser praticados e assinados pelo servidor judiciário, prescindindo de determinação expressa do juízo”.⁷⁴²

Da redação do art. 203, §4º, do Código de Processo Civil extrai-se que os atos ordinatórios “independem de despacho, devendo ser praticados de ofício pelo servidor e revistos pelo juiz quando necessário”, tal regra implica na implementação dos princípios do impulso oficial e da razoável duração do processo (previstos, respectivamente, no art. 2º e no art. 4º do Código de Processo Civil). Os referidos atos podem ser classificados em atos de movimentação, documentação, execução e comunicação processual⁷⁴³, sendo que após o advento do processo eletrônico os mesmos passaram a ser realizados, em grande parte, mediante sistemas computadorizados, na forma autorizada pelos artigos 4º a 9º da Lei 11.419/2006. A referida realidade já está sedimentada na prática jurídica brasileira.

<https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=11043891>. Acesso em: 12 jan. 2023.

⁷⁴² “A uma, porque os atos meramente ordinatórios, normalmente praticados por meio de despachos, não possuem a potencialidade de causar prejuízo a qualquer das partes envolvidas nos processos, tanto que podem ser delegados, nos termos do art. 93, XIV, da Constituição Federal” (...)

“Na verdade, as normas processuais já autorizam expressamente o servidor a proferir atos ordinatórios, reputados como necessários à tramitação regular do processo, como por exemplo, conceder vistas às partes, juntar petições, expedir ofícios etc. Sabe-se que esses atos administrativos e de expediente facilitam a celeridade do processo e podem ser praticados e assinados pelo servidor judiciário, prescindindo de determinação expressa do juízo”.

(BRASIL. Superior Tribunal de Justiça (6. Turma). **Recurso Especial 1177612/SP**. *Habeas corpus* concedido de ofício em razão da não configuração da tipicidade subjetiva para o crime de falsificação de documentos públicos (decisões judiciais). Brasília, DF: Superior Tribunal de Justiça, [2011]. Relator: Min. Og Fernandes, 01 de setembro de 2011. Disponível em: https://scon.stj.jus.br/SCON/GetInteiroTeorDoAcordao?num_registro=201000156359&dt_publicacao=17/10/2011. Acesso em: 12 jan. 2023).

⁷⁴³ “A movimentação e a documentação fazem-se precipuamente através do escrivão e de seus funcionários (escreventes). São atos de movimentação processual: a conclusão dos autos ao juiz, a vista às partes, a remessa ao contador, a expedição de mandados e ofícios. São atos de documentação: a lavratura dos termos referentes à movimentação (conclusão, vista etc.), a feitura do termo de audiência, o lançamento de certidões, etc.

A execução é ordinariamente um encargo do oficial de justiça: trata-se de atos realizados fora dos auditórios e cartórios, em cumprimento a mandado judicial (penhora, prisão, busca-e-apreensão etc.).

A comunicação processual, consistente em citações ou intimações, é realizada pelo escrivão, com auxílio dos Correios, ou pelo oficial de justiça, em cumprimento de mandados judiciais”.

(CINTRA, Antônio Carlos de Araújo; GRINOVER, Ada Pellegrini; DINAMARCO, Cândido Rangel. Op. cit., p. 358).

Nessa linha, admite-se a delegação da realização de atos processuais aos servidores do Judiciário, havendo legalidade e legitimidade na realização pelos servidores de atos processuais “não decisórios”. Percebe-se que a mesma delegação realizada aos servidores pode implicar na delegação da execução de atos de mero expediente para sistemas computacionais (com ou sem Inteligência Artificial).

Por outro lado, a distinção entre ato de mero expediente (expressamente delegável) e despachos (art. 203, §3º, do CPC) aparentemente pode ser um entrave à delegação de atos processuais, a qual é inerente à teoria processual, uma vez que os sistemas computacionais baseados em Inteligência Artificial cada vez mais são desenhados para atividades que decorrem da atividade decisória (penhora, uso de sistemas de consultas, quebras de sigilos bancário, fiscal e de dados, etc.).

Nessa linha, a regra do referido art. 203, §3º, do CPC estabelece que são “despachos todos os demais pronunciamentos do juiz praticados no processo, de ofício ou a requerimento da parte”, em relação aos quais “não cabe recurso” (art. 1.001 do CPC), uma vez que **não possuem a “natureza decisória”** expressamente atribuída às decisões interlocutórias e às sentenças (art. 203, §1º e §2º, do CPC). Dessa forma, os “despachos de mero expediente, são pronunciamentos judiciais **sem qualquer conteúdo decisório** e destinados à boa organização do processo, com seu encaminhamento rumo ao ato final que será a sentença” [destaque não original]⁷⁴⁴.

Logo, na forma do desafio “mais-por-menos” definido por Richard Susskind⁷⁴⁵, podem ser automatizadas, mediante Inteligência Artificial, até mesmo as movimentações processuais que, numa abordagem processual tradicional ainda alheia à Inteligência Artificial⁷⁴⁶, dependeriam de atuação da figura do juiz mediante a elaboração de despachos, haja vista a baixa densidade decisória dos despachos de mero expediente.⁷⁴⁷ Numa leitura de interpretação sistemática da Constituição Federal

⁷⁴⁴ DINAMARCO, Cândido Rangel; LOPES, Bruno Vasconcelos Carrilho. **Teoria Geral do Novo Processo Civil**. 2ª ed. São Paulo: Malheiros, 2017, p. 186-187.

⁷⁴⁵ “*I believe the more-for-less challenge will underpin the next decade of legal service. The more-for-less challenge will, I expect, irreversibly change the way that lawyers work*” (SUSSKIND, Richard. 2017, p. 5).

⁷⁴⁶ “*Limitando el análisis a las justicias civil y penal, existen diversos aspectos que actualmente parecen complicados porque responden a variables y situaciones del pasado, pero que ya han perdido vigencia, aunque los legisladores no hayan reaccionado todavía en su mayoría. Por ello, parece razonable comenzar con un breve compendio actualizador de ambos procesos, poniéndolos en el contexto actual, labor que quizá se hubiera podido hacer, no obstante, hace ya algún tiempo*” (FENOLL, Jordi Nieva. **Inteligencia Artificial y Proceso Judicial**. Madri: Marcial Pons, 2018, p. 33).

⁷⁴⁷ “*Y siendo así, las demandas deberían ser admitidas automáticamente, siendo entregadas a la contraria inmediatamente para su contestación. Solamente la solicitud de medidas cautelares haría precisa la intervención humana en esta fase de alegaciones. Y tras la contestación, de nuevo esa*

(pautada na unidade da Constituição⁷⁴⁸) e das demais normas aplicáveis, a atividade decisória em sentido estrito (sentenças e decisões interlocutórias) resta intocada sob o domo do dever de independência funcional do magistrado (art. 93 da Constituição Federal, art. 35, I, da Lei Complementar 35/1979 e do art. 4º ao art. 7º do Código de Ética da Magistratura⁷⁴⁹), o que possibilita o uso da IA processual.

A partir das premissas aqui lançadas, evidencia-se que existe uma série de atos processuais que podem ser executados mediante algoritmos de lógica simples (no modelo *if - then*⁷⁵⁰) ou mediante *machine learning*⁷⁵¹, a fim de que o processo tenha um andamento mais rápido. Assim, emerge uma era⁷⁵² (da quarta revolução industrial) na qual o ateliê de andamento processual (decorrente do entrelaçamento entre cartório⁷⁵³ e gabinete do Juízo), pautado em atividade humana (de servidores, estagiários e magistrados), terá seu funcionamento de maneira mais fluida e pervasiva⁷⁵⁴ em decorrência da Inteligência Artificial.

Nesse aspecto, a presente abordagem coloca-se em uma distância cautelosa do pesadelo dataísta referido por Yuval Harari⁷⁵⁵, na medida em que a robotização

intervención del juez solamente se haría precisa si se formulan excepciones procesales. Y hay que tener en cuenta que muchas de las mismas, las que pueden resol verse actualmente en una audiencia previa, también son automatizables. La falta de capacidad o de representacion depende de datos que constan en registros oficiales, tales como la menor edad, la representación de sociedades de capital o las incapacitaciones. Lo mismo sucede con la cosa juzgada o la litispendencia, pues requiriéndose en esta fase prematura una perfecta identidad de los objetos de los juicios para ser apreciada, la misma podría ser también automatizada. Algo parecido podría suceder con las situaciones de litisconsorcio necesario, si la jurisprudencia concretara con auténtica precisión los casos en que se produce. Y también podría acontecer algo similar con la inadecuación de procedimiento, puesto que basta con que la máquina coteje lo solicitado con el procedimiento escogido, así como con la demanda defectuosa. Y es que, como ya se dijo, actualmente las máquinas ya interpretan los textos que 'leen'. Si la demanda es inteligible, el sistema puede rechazar la excepción automáticamente, haciendo lo opuesto en caso contrario. Igual que podría rechazar una acumulación improcedente de pretensiones, al menos en los casos más obvios" (Ibid., p. 34-35).

⁷⁴⁸ MÜLLER, Friedrich. **Metodologia do Direito Constitucional**. 4.^a ed. rev. e atual. e ampl. Tradução: Peter Naumann. São Paulo: Revista dos Tribunais, 2010, p. 82.

⁷⁴⁹ CONSELHO NACIONAL DE JUSTIÇA. **Resolução nº 60, de 19 de setembro de 2008**.

⁷⁵⁰ VEGA, Italo S. Inteligência Artificial e Tomada de Decisão – A necessidade de agentes externos. In: FRAZÃO, Ana (coord.) et. al. **Inteligência Artificial e Direito: Ética, Regulação e Responsabilidade**. São Paulo: Thomson Reuters Brasil, 2019, p. 103.

⁷⁵¹ MEDVEDEVA, Masha; VOLS, Michel; WIELING, Martijn. Using machine learning to predict decisions of the European Court of Human Rights. **Artificial Intelligence and Law**, Dordrecht, 2019. DOI: 10.1007/s10506-019-09255-y.

⁷⁵² HARTMANN PEIXOTO, Fabiano; SILVA, Roberta Zumblick Martins da. Op. cit., p. 70-71.

⁷⁵³ Em alguns Tribunais denominado como Secretaria.

⁷⁵⁴ "By its very nature, a computing so pervasive and so deeply intertwined with everyday life will exert a transformative influence on our relationships with ourselves and with each other" (GREENFIELD, Adam. Op. cit., p. 84).

⁷⁵⁵ "O dataísmo reúne os dois, assinalando que exatamente as mesmas leis matemáticas se aplicam tanto aos algoritmos bioquímicos como aos eletrônicos. O dataísmo, portanto, faz ruir a barreira entre animais e máquinas com a experiência de que, eventualmente, os algoritmos eletrônicos decifrem e superem os algoritmos bioquímicos" (HARARI, Yuval Noah. **Homo Deus**, p. 370).

processual⁷⁵⁶ aqui pensada se volta às atividades padronizadas e repetitivas, justamente para liberar os servidores do Judiciário e os magistrados para a realização de funções criativas inerentes ao pensamento humano.⁷⁵⁷

Por outro lado, a utilização de Inteligência Artificial processual (inserida dentro do modelo *Big Data*⁷⁵⁸) é um risco⁷⁵⁹ que deve ser sempre levando em conta no desenvolvimento dos sistemas, conforme foi tratado na Lei Geral de Proteção de Dados (Lei 13.709/2018) e sempre deve nortear a pesquisa (privada, governamental e acadêmica⁷⁶⁰), para que não haja mau uso dos dados obtidos ou lesão a direitos dos cidadãos.

Portanto, numa visão constitucionalizada do processo e de sua evolução tecnológica (o que demanda leitura sistemática da Constituição Federal e interpretação evolutiva⁷⁶¹), existe a possibilidade de delegação à IA (art. 93, XVI, da Constituição Federal) de tarefas vinculadas aos atos processuais de maior simplicidade (pautados em atos ordinatórios e despachos de mero expediente – art. 203, §3º e §4º, do CPC), as quais antes eram reservadas apenas aos seres humanos (servidores e juízes).

É bastante conhecido, nos campos militar e da política internacional, o fato de que a África do Sul, após o gasto de muitos recursos e bastante tempo em pesquisa científica, supostamente decidiu “esquecer” como se faz uma bomba nuclear após já possuir tal tipo de armamento⁷⁶². Ainda que isso levante dúvidas quanto à veracidade do fato alardeado pelos sucessivos governos daquele país, o conceito é bastante interessante. O esquecimento deliberado como forma de proteção.

Nesse sentido, o estabelecimento de regras padrão para decisões é a forma mais conhecida de desenvolvimento do direito, conforme se percebe do pensamento jurídico desde Herbert Hart (e seu conceito de poder discricionário regulamentar

⁷⁵⁶ FENOLL, Jordi Nieva. Op. cit., p. 34.

⁷⁵⁷ KAHNEMAN, Daniel. Op. cit., posição 281.

⁷⁵⁸ MAGRANI, Eduardo. **Entre Dados e Robôs**, p. 200.

⁷⁵⁹ O'NEIL, Cathy. Op. cit., p. 204.

⁷⁶⁰ HARTMANN PEIXOTO, Fabiano. **Inteligência Artificial e Direito: Convergência Ética e Estratégica**, p. 25.

⁷⁶¹ BARROSO, Luís Roberto. **Interpretação e Aplicação da Constituição**. 3ª ed. São Paulo: Saraiva, 1999, p. 124.

⁷⁶² VAN WYK, Jo-Ansie. South Africa's Nuclear Diplomacy Since The Termination of Its Nuclear Weapons Programme. **Scientia Militaria - South African Journal of Military Studies**, v. 42, p. 80-101. DOI: 10.5787/42-1-1082.

atribuído ao Judiciário para complementação da norma)⁷⁶³ até Jürgen Habermas.⁷⁶⁴ Há um campo de legitimidade decisória nos trâmites processuais que sempre restará (e incumbirá) à figura do juiz, a qual tem a possibilidade, desde que respeitadas as regras processuais de ordem pública, de moldar o andamento do processo da forma que melhor atenda à eficiência, à razoável duração do processo e à celeridade⁷⁶⁵ (sem descuidar das garantias do devido processo legal, do contraditório e da ampla defesa⁷⁶⁶).

Conforme referido no art. 93, XIV, da Constituição Federal⁷⁶⁷, **cada juiz, no âmbito de sua independência funcional**⁷⁶⁸ (e não qualquer órgão administrativo de tribunal⁷⁶⁹), pode adotar padrão especialmente desenhado⁷⁷⁰ de definição de regras mediante **delegação de atos de maior simplicidade** (inclusive os relacionados a despachos), de modo a criar regras padrão a serem seguidas pelos algoritmos de Inteligência Artificial (e pelos servidores da unidade jurisdicional), a fim de garantir a automatização do trâmite processual.

As portarias e as ordens de serviço redigidas pelos magistrados, desde que utilizadas com racionalidade, têm justamente o importante papel de definir a **regra padrão de trâmite do processo**, pois “uma regra padrão permite que as pessoas

⁷⁶³ HART, Herbert L. A. Op. cit., p. 171.

⁷⁶⁴ “Para os sistemas jurídicos modernos, recomenda-se uma dupla delimitação. Ao direito enquanto sistema de ação, podemos subordinar a totalidade as interações reguladas através das normas”. (...) “Ora, o sistema de direito em sentido mais estrito ultrapassa essa concepção. Ele abrange todas as interações, também as quais que não se orientam pelo direito, podendo produzir direito novo e reproduzi-lo enquanto tal. Para a institucionalização do sistema jurídico neste sentido, necessita-se da auto-aplicação do direito através de regras secundárias que constituem e transmitem as competências da normatização, da aplicação e da imposição do direito. Os ‘poderes’ do Estado, da legislação, da justiça e da administração distinguem-se de acordo com essas funções” (HABERMAS, Jürgen. **Direito e Democracia**, p. 242).

⁷⁶⁵ Art. 5º, LXXVIII, e art. 37, caput, da Constituição Federal e art. 4º e art. 8º do Código de Processo Civil.

⁷⁶⁶ Art. 5º, LIV e LV, da Constituição Federal e art. 7º e art. 8º do Código de Processo Civil.

⁷⁶⁷ “Art. 93. (...)

(...)

XIV - os servidores receberão delegação para a prática de atos de administração e atos de mero expediente sem caráter decisório” (BRASIL. **Emenda Constitucional nº 45, de 30 de dezembro de 2004**. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/emendas/emc/emc45.htm. Acesso em: 12 jun. 2022).

⁷⁶⁸ Art. 5º, XXXV, e art. 93, caput, da Constituição Federal, art. 35, I, da Lei Complementar 35/1979 (Lei Orgânica da Magistratura) e art. 4º e art. 7º do Código de Ética da Magistratura.

⁷⁶⁹ Seja por meio de seu Pleno, do Órgão Especial, da Presidência ou da Corregedoria.

⁷⁷⁰ Neste aspecto o art. 5º do Código de Ética da Magistratura é bastante claro ao prever que “Impõe-se ao magistrado pautar-se no desempenho de suas atividades sem receber indevidas influências externas e estranhas à justa convicção que deve formar para a solução dos casos que lhe sejam submetidos” (CONSELHO NACIONAL DE JUSTIÇA. **Resolução nº 60, de 19 de setembro de 2008**).

escolham escolher - e também escolham não escolher (confiando no padrão)” [tradução livre]⁷⁷¹.

A escolha do magistrado ao lançar mão de normativas internas da unidade jurisdicional é heurística⁷⁷² voltada para o desestímulo à escolha de opções menos eficientes (ou mais questionáveis do ponto de vista da legitimidade jurídica) por parte de servidores que atuam no processo e dos sistemas de Inteligência Artificial. Assim, tem-se o estabelecimento de um padrão baseado em heurística da disponibilidade⁷⁷³, uma vez que as “regras padrão geralmente levam à automação para ajudar a superar vários vieses comportamentais e também para responder ao fato de que os seres humanos inevitavelmente têm uma ‘largura de banda’ limitada” (tradução livre)⁷⁷⁴.

O referido esquecimento funcional é uma manifestação da própria heurística da disponibilidade⁷⁷⁵, uma vez que na “dúvida, o Sistema 1 aposta numa solução, e a aposta é orientada pela experiência”.⁷⁷⁶ Ressalta-se, aqui, a divisão do pensamento humano entre o intuitivo e rápido, Sistema 1, e o reflexivo e lento, Sistema 2⁷⁷⁷, na forma desenvolvida por Daniel Kahneman. A função do Sistema 1 é manter (e atualizar) o seu modelo de mundo (que representa o que existe nele), o qual “é construído por associações que ligam ideias de circunstâncias, eventos, ações e resultados que coocorrem com alguma regularidade, seja ao mesmo tempo, seja dentro de um intervalo relativamente curto”⁷⁷⁸.

Vê-se que a intuitividade da grande maioria dos andamentos processuais é causada pela simplicidade das opções de escolha que podem ser feitas em relação ao trâmite dos processos, o que contribuí sensivelmente para a redução do risco de vieses de representação⁷⁷⁹ ou de disponibilidade⁷⁸⁰, haja vista o rarefeito caráter decisório.

⁷⁷¹ SUNSTEIN, Cass R. **Choosing Not to Choose**, p. 9.

⁷⁷² “Se você acabar emperrado num cenário intratável, lembre-se de que heurística, aproximações e o uso estratégico da aleatoriedade podem ajudá-lo a encontrar soluções viáveis” (CHRISTIAN, Brian; GRIFFITHS, Tom. **Algoritmos para viver**, posição 6358-6359).

⁷⁷³ KAHNEMAN, Daniel. Op. cit., posição 461-463.

⁷⁷⁴ SUNSTEIN, Cass R. **Choosing Not to Choose**, p. 14.

⁷⁷⁵ “Definimos a heurística de disponibilidade como o processo de julgar a frequência segundo a ‘facilidade com que as ocorrências vêm à mente’” (KAHNEMAN, Daniel. Op. cit., posição 2652-2654).

⁷⁷⁶ “As regras da aposta são inteligentes: eventos recentes e o contexto presente têm o maior peso em determinar uma interpretação” (Ibid., posição 1675).

⁷⁷⁷ Ibid., posição 281.

⁷⁷⁸ Ibid., posição 1502.

⁷⁷⁹ Ibid., posição 490-499.

⁷⁸⁰ Ibid., posição 477-478.

O estudo das heurísticas e vieses é um campo de enorme riqueza e é bastante promissor no direito. Todavia, a implementação de algoritmos de Inteligência Artificial nos sistemas de processo eletrônico (por meio do estabelecimento de decisões padrão em normativas internas de cada unidade jurisdicional) tem o potencial de combater o que se pode denominar **viés do estagiário** (com todo o respeito à classe da qual todos os operadores jurídicos já fizeram parte), o qual pode ser classificado como uma expressão da repetição impensada de padrões (causada por uma mistura de insegurança, pouco treinamento e ausência de conhecimento especializado) e que, em certas condições, pode inclusive variar para a mudança repentina de comportamento.

Além disso, o **viés do especialista** (referido por Nassim Taleb⁷⁸¹) também deve ser minorado na implementação de IA mediante regras padrão, pois no caso de evento de grande imprevisibilidade, ou do acúmulo marginal de eventos que levem à situação imprevista, haverá a necessidade da adoção de uma arquitetura de decisões em que os algoritmos simplesmente deixem de aplicar o modelo padrão de decisão e demandem supervisão humana de maneira ativa (e isso passa pelo debate ético dos sistemas de IA⁷⁸²), pois os padrões decisórios de andamentos processuais são feitos para situações de normalidade, não para excepcionalidades que devem ser objeto de tratamento pelo magistrado e pelos servidores da unidade jurisdicional.

A partir de tais premissas, há ponto de partida para se instrumentalizar o uso de Inteligência Artificial⁷⁸³ processual, de modo que os trâmites processuais inerentes aos atos de maior simplicidade possam ser executados de maneira automática ou semiautomatizada, o que traz a função jurisdicional para a realidade da era da IA.

Ainda que o robô processual (não tão questionador quanto os robôs de Asimov) esteja distante de realizar a introspecção tipicamente humana, percebe-se que há um

⁷⁸¹ “Não se pode ignorar a autoilusão. O problema com especialistas é que eles não sabem o que não sabem. Falta de conhecimento e ilusão quanto à qualidade do conhecimento coincidem – o mesmo processo que faz com que você saiba menos também o deixa satisfeito com o seu conhecimento” (TALEB, Nassim N. Op. cit., p. 195).

⁷⁸² YAPO, Adrienne; WEISS, Joseph. Op. cit.

⁷⁸³ “Há um bom tempo que os fluxos de tarefas administrativas, tanto para direcionamento do processo judicial em si quanto para desenvolver serviços de apoio à atividade jurisdicional tomam grande quantidade de esforços, tempo e recursos e recursos e outros elementos precisos ao serviço público. Também não é de hoje que esses fluxos estão a cada momento mais precisos e produzem muitos elementos de métricas de eficiência e eficácia. Há, portanto, um bom substrato para aplicação e eficácia dos em IA homologada e implementada” (HARTMANN PEIXOTO, Fabiano; SILVA, Roberta Zumblick Martins da. Op. cit., p. 120).

mínimo de habilidades que já podem ser exigidas de sistemas de Inteligência Artificial processual.

Evidencia-se que o desenvolvimento multiplataforma da Inteligência Artificial processual (pelos tribunais, pela sociedade, pela academia e pelo mercado) é um caminho democrático, garantido pelo art. 8º da Lei 11.419/2006, cabendo ao Conselho Nacional de Justiça a atuação regulatória (legítima⁷⁸⁴ e constitucional⁷⁸⁵) quanto à qualidade e à operabilidade do sistema.

No modelo de processo instrumentalizado pela Inteligência Artificial perde funcionalidade a distinção forte entre atos ordinatórios e despachos de mero expediente, uma vez que ambos os tipos de atos processuais mais simples referidos são o objeto de melhor destaque para a automatização.

O robô processual nos moldes aqui pensados é uma ideia possível de ser executada do ponto de vista jurídico, pois se pauta na busca de automatização de tarefas repetitivas por meio de Inteligência Artificial, vindo a prestar auxílio aos servidores do Judiciário e aos juízes numa prática de trâmites processuais que utilize as soluções algorítmicas já atualmente⁷⁸⁶ passíveis de serem desenvolvidas e implementadas.

Dessa forma, ainda que temas como a singularidade e uma Inteligência Artificial que consiga auxiliar os juízes nas decisões judiciais sejam alguns dos mais sensíveis, a espera pela solução das grandes questões da relação entre IA e direito não pode

⁷⁸⁴ “O CNJ vem afirmando a sua importância institucional ao defender o papel do Judiciário como estabilizador de expectativas econômicas e garantidor de direitos fundamentais, bem como a relevância da cooperação com outras instituições públicas e com a iniciativa privada para a construção de um Judiciário mais ajustado às demandas da sociedade” (MENDES, Gilmar Ferreira. *Parcerias Institucionais e o Conselho Nacional de Justiça: instrumento de efetivação do acesso à Justiça*. **Revista Jurídica da Presidência**, Brasília, v. 20, n. 122, p. 516-540, Out. 2018/Jan. 2019. DOI: 10.20499/2236-3645.RJP2018v20e122-1819).

⁷⁸⁵ “A construção do plano estratégico dos tribunais representa a forma como o CNJ encontrou para padronizar importantes questões administrativas, financeiras e orçamentárias das justiças federal e estaduais, equilibrando com a autonomia dos tribunais e das justiças estaduais. O CNJ, órgão centralizador e federal, apresenta a estrutura do plano estratégico e os indicadores do plano.

Os tribunais, por sua vez, exercendo as autonomias administrativa, financeira e orçamentária estabelecem seus planos. De maneira conjunta, CNJ e tribunais fixam as metas que devem ser buscadas por todos os tribunais, as quais são fiscalizadas pelos tribunais e pelo CNJ. Essa foi a maneira estabelecida pelo federalismo judicial brasileiro para equilibrar centralização e descentralização na busca por *accountabilities* judiciais comportamental e institucional” (ROBL FILHO, Ilton Norberto. **Accountability e Independência Judiciais: O Desenho Institucional do Judiciário e do Conselho Nacional de Justiça no Estado Democrático de Direito Brasileiro**. 2012. Tese (Doutorado em Direito). Universidade Federal do Paraná (UFPR), Curitiba, 2012, p. 228).

⁷⁸⁶ CRUZ, Fabricio Bittencourt da; BRITTO, Melina Carla de Souza; MOREIRA, Guilherme Martelli; BRITTO JUNIOR, Alceu de Souza. Robôs Substituem Juízes? O Estado da Arte da Inteligência Artificial no Judiciário Brasileiro. **Revista Antinomias**, v. 3, n. 1, p. 8-41, jan./jun., 2022.

obstar a aplicação de soluções já existentes para os problemas mais simples e de mais fácil trato de pelos sistemas de computador.

Assim, desviar a capacidade cognitiva de seres humanos para a realização de tarefas repetitivas deve ser cada vez menos comum, a fim de que o processo eletrônico ganhe novas potencialidades com a utilização da Inteligência Artificial.

Nesse caminho, a delegação às máquinas de rotinas (antes reservadas aos seres humanos mediante atos ordinatórios e despachos de mero expediente) pauta-se na segurança e na celeridade dos trâmites, o que pode ser garantido por uma arquitetura de decisões que espelhe heurísticas desenhadas pelo próprio julgador (conforme sua cultura jurídica e sua compreensão do processo), desde que respeitados os demais aspectos da legislação processual e os direitos fundamentais dos jurisdicionados.

O robô processual, dentro de sua simplicidade conceitual voltada à automatização de atos de mero expediente e atos que decorreriam de despachos de mero trâmite do processo, ainda que mexa com antigos modelos de funcionamento do processo, é uma fase inicial da transcendência do processo judicial para a nova era e para sua nova forma, na qual haverá o auxílio de sistemas de Inteligência Artificial judicial às atividades processuais humanas.

Sendo assim, diferentemente de decisões interlocutórias e sentenças, nas quais se desenha a utilização de sistemas de Inteligência Artificial voltados à cooperação com o julgador humano na busca da decisão correta, os sistemas de IA judicial voltados ao mero impulsionamento do processo (trâmite passo a passo do processo, intimações das partes, elaboração de documentos simples, formalização de atos processuais, cálculos judiciais com premissas previamente estabelecidas, remessa do processo ao gabinete do magistrado, etc.) podem ter um grau maior de autonomia, mediante a definição prévia de cada Tribunal e a escolha final de cada magistrado em sua respectiva unidade judicial. Por exemplo, o ato de se intimar uma parte porque um novo documento foi apresentado no processo (pela parte adversa, por um perito ou por qualquer outra pessoa) apresenta baixo grau de complexidade e poucos riscos quando se compara à decisão judicial que determina penhora eletrônica em contas bancárias, à análise probatória ou à dosimetria da pena.

Em suma, ao se desenvolver sistemas de IA judicial, deve existir razoabilidade na aplicação das regras limitativas de seu desenvolvimento, de maneira que sistemas

de simples automatização de andamentos processuais (robôs processuais) tenham um tratamento adequado, e, menos restritivo, do que sistemas de auxílio à fundamentação de decisões judiciais. São problemas diversos em complexidade e riscos, com soluções de IA judicial proporcionalmente diversas, sendo inviável se igualar tais casos num plano das técnicas de IA disponíveis ou mesmo do peso da aplicação das travas regulatórias.

3.3 Inteligência Artificial Judicial: a Pluralidade Epistêmica e o Juiz da Era da IA

A figura do juiz exerce inafastável função constitucional (art. 5º, XXXV, da Constituição Federal), consistente no exercício da atividade jurisdicional, o que decorre como um efeito natural da separação harmônica de Poderes prevista no art. 2º, *caput*, da Constituição Federal.

O ato de julgar consiste na tomada de decisão através de pronunciamentos a respeito de questões trazidas pelas partes, cujas espécies são sentença, decisão interlocutória e despacho (art. 203 do Código de Processo Civil).

No exercício da atividade julgadora a tomada de decisão se desenvolve em dois planos, primeiramente no plano mental do julgador⁷⁸⁷ (o qual não pode ser fotografado, filmado ou traduzido em dados) para, apenas depois, se traduzir num processo justificatório⁷⁸⁸ analítico⁷⁸⁹ dos motivos da decisão, justificação esta que pode tomar variadas formas, ainda que muitas vezes os resultados interpretativos sejam os mesmos.

As teorias da argumentação jurídica representaram processo lógico-interpretativo para a tomada de decisão judicial⁷⁹⁰, num primeiro aspecto, e servem como manto de racionalidade institucional às decisões tomadas⁷⁹¹, sendo que as

⁷⁸⁷ “A figuração lógica dos fatos é o pensamento” (WITTGENSTEIN, Ludwig. **Tractatus Logico-Philosophicus**. Tradução: Luiz Henrique Lopes dos Santos. São Paulo: Editora Universidade de São Paulo, 2020, p. 139).

sua decisão no funcionamento interno do poder Judiciário.

⁷⁸⁸ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 162.

⁷⁸⁹ Na forma dos elementos do art. 489, §1º, do Código de Processo Civil.

⁷⁹⁰ “(...) a teoria da argumentação ganha significado especial, porque é dela a tarefa de reconstruir o pressupostos e condições formal-pragmáticos de um comportamento explicitamente racional” (HABERMAS, Jürgen. **Teoria do Agir Comunicativo**, 1, p. 21).

⁷⁹¹ MACCORMICK, Neil. **Retórica y Estado de Derecho**: Una teoría del razonamiento jurídico. Lima: Palestra Editores, 2017, posição 490. E-book Kindle.

mesmas têm margem de absorção de outros modelos teóricos⁷⁹² sem que isso implique em exclusão dos modelos argumentativos, pois os padrões que as teorias da argumentação utilizam são um caminho de explicabilidade⁷⁹³ racional para as decisões judiciais auxiliadas por Inteligência Artificial. Nessa linha, o pensamento necessário para a estruturação de argumentações racionais se assemelha com a programação, a qual “significa a capacidade de formular problemas, pensar criativamente em soluções e expressar uma solução de forma clara e precisa”.⁷⁹⁴

Dessa forma, no cotidiano da atividade jurisdicional não há uma determinação (e nem há uma prática recorrente) de que os juízes (em sua maioria) pensem em Teoria do Direito no momento de decidir⁷⁹⁵, haja vista que os julgadores estão envolvidos numa cultura de normatividade, que é atrelada à interpretação de diplomas legais, ao uso de jurisprudência e à aplicação de lições doutrinárias mais ligadas às especificidades de cada ramo do direito. Isso ocorre porque o juiz não necessita ser um teórico do Direito para o exercício de sua atividade.⁷⁹⁶ Ademais, a legislação brasileira traça diretrizes consideravelmente abertas acerca de um modelo de precedentes (no Código de Processo Civil), da observância às Súmulas Vinculantes (art. 103-A da Constituição Federal), de métodos de integração do ordenamento jurídico (art. 4º do Decreto-Lei 4.657/1942⁷⁹⁷) e uma abordagem consequencialista típica de razão prática (art. 20 do Decreto-Lei 4.657/1942⁷⁹⁸). Dessa maneira, os limites às espécies de argumentos que são considerados válidos numa decisão judicial são relativamente abertos no plano epistêmico-jurídico, ainda que tenha

⁷⁹² Por exemplo, o posicionamento de mútua cooperação e complementariedade entre hermenêutica e argumentação em: MARRAFON, Marco Aurélio. Op. cit., p. 99-106.

⁷⁹³ DEEKS, Ashley. Op. cit.

⁷⁹⁴ DOWNEY, Allen B. **Pense em Python**: pense como um cientista da computação. Tradução: Sheila Gomes. São Paulo: Novatec, 2016, p. 25.

⁷⁹⁵ Embora o cotidiano da atividade judicial não envolva pensar sobre as Teorias do Direito e seus diferentes reflexos sobre as decisões, em casos especiais pode existir debate sobre o que é o Direito, quais são os seus elementos e como as diferentes concepções jurídicas se aplicam no julgamento. Exemplos disso são julgamentos dos casos *Elmer (Riggs vs. Palmer)* e *snail darter (Tennessee Valley Authority vs. Hill)* citados por Dworkin ao problematizar sua concepção de Direito (DWORKIN, Ronald. **O Império do Direito**. 3ª ed., p. 20-29).

⁷⁹⁶ Ainda assim, quando um Tribunal adota a interpretação que entende correta para um determinado tema, ele define o sentido do Direito para caso sob julgamento e para as pessoas alcançadas pela decisão.

⁷⁹⁷ Lei de Introdução às Normas do Direito Brasileiro (LINDB).

⁷⁹⁸ “Art. 20. Nas esferas administrativa, controladora e judicial, não se decidirá com base em valores jurídicos abstratos sem que sejam consideradas as consequências práticas da decisão. Parágrafo único. A motivação demonstrará a necessidade e a adequação da medida imposta ou da invalidação de ato, contrato, ajuste, processo ou norma administrativa, inclusive em face das possíveis alternativas” (BRASIL. **Lei nº 13.655, de 25 de abril de 2018**. Disponível em: https://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2018/Lei/L13655.htm Acesso em: 11 jun. 2022).

existido o referido esforço normativo para reduzir os argumentos passíveis de utilização pelos julgadores àqueles que são institucionalmente aceitos.

Apenas como um exemplo dessa relativa abertura do sistema para decisões diferentes, num hipotético caso de demissão coletiva, ao decidir pela reintegração de trabalhadores demitidos um Juiz do Trabalho pode estar preocupado com a preservação de empregos, uma heurística constitucional de matriz típica do Direito do Trabalho e do Direito Processual do Trabalho. Enquanto, em caso análogo, um Juiz de Vara de Falências pode vir a se preocupar com a liberdade de iniciativa e com a preservação de empresa que se encontra em situação financeira delicada, princípios mais ligados ao Direito Empresarial (embora igualmente alcancem fundamento na Constituição). E neste exemplo, a depender da Teoria do Direito que o juiz utilize (Alexy⁷⁹⁹, Dworkin⁸⁰⁰ ou MacCormick⁸⁰¹, por exemplo) ocorrerá uma variação no resultado do julgamento.

Noutro exemplo, antes do tema ter sido tratado por meio do instituto da Repercussão Geral⁸⁰², nos casos que envolvem a aplicação de convenções internacionais aos contratos de transporte aéreo internacional o Superior Tribunal de Justiça tinha como preocupação garantir a proteção do consumidor nessa espécie de relação jurídica, o que levava o Superior Tribunal de Justiça a utilizar a teoria do diálogo das fontes (reforçada, no fundo, pela substância constitucional da proteção ao consumidor⁸⁰³) para alegar que o Código de Defesa do Consumidor seria superior à regra do art. 178 da Constituição Federal.⁸⁰⁴ Já o Supremo Tribunal Federal afasta a aplicação do Código de Defesa do Consumidor para resguardar a aplicação da

⁷⁹⁹ Que iguala os princípios no plano abstrato e em potencial de aplicabilidade em sua teoria da colisão (ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 94-103).

⁸⁰⁰ Que distingue princípios (*principles*) e políticas (*policies*) já no plano abstrato (DWORKIN, Ronald. **Levando os Direitos a Sério**, p. 36-50).

⁸⁰¹ Que compreende aspectos consequencialistas de determinados posicionamentos judiciais (MACCORMICK. **Argumentação Jurídica e Teoria do Direito**, p. 145-152).

⁸⁰² Repercussão Geral de Tema 210 do Supremo Tribunal Federal.

⁸⁰³ Art. 5º, XXXII, e art. 170, V, da Constituição Federal.

⁸⁰⁴ “A responsabilidade civil das companhias aéreas em decorrência da má prestação de serviços, após a entrada em vigor da Lei n. 8.078/90, não é mais regulada pela Convenção de Varsóvia e suas posteriores modificações (Convenção de Haia e Convenção de Montreal) ou pelo Código Brasileiro de Aeronáutica, subordinando-se ao Código de Defesa do Consumidor” (BRASIL. Superior Tribunal de Justiça (3. Turma). Agravo Regimental nos Embargos de Declaração no Agravo em Recurso Especial 418.785/RJ. Aplicação do Código de Defesa do Consumidor aos casos de responsabilidade civil decorrente de transporte aéreo internacional em detrimento das normas da Convenção de Haia e Convenção de Montreal. Brasília, DF: Superior Tribunal de Justiça, [2016]. Relator: Min. João Otávio de Noronha, 17 de maio de 2016. Disponível em: https://scon.stj.jus.br/SCON/GetInteiroTeorDoAcordao?num_registro=201303584762&dt_publicacao=23/05/2016. Acesso em: 09 jul. 2022).

referida norma da Constituição de 1988⁸⁰⁵. Ainda que se discorde da linha de precedentes do Superior Tribunal de Justiça anterior à decisão vinculante do Supremo Tribunal Federal, havia, naquele momento, um posicionamento resguardado do ponto de vista institucional, mesmo que tenha sido combatido e superado parcialmente pela Corte constitucional.⁸⁰⁶ Esse último exemplo demonstra como, mesmo com argumentos racionais e elementos normativos bem alicerçados, no caso da regra do art. 178 da Constituição Federal, o uso de teorias interpretativas diferentes leva a resultados diferentes nos julgamentos. A partir disso, constata-se que há uma pluralidade epistêmica que, respeitados os limites institucionais democráticos, pode ser exercida pelos magistrados e pelos Tribunais em suas decisões.

Em resumo, há a possibilidade, no âmbito acadêmico e na prática do Direito, de se discordar, por exemplo, da teoria do diálogo das fontes⁸⁰⁷, da hermenêutica filosófica⁸⁰⁸ ou da teoria positivista normativista de Kelsen⁸⁰⁹, mas a discordância com o modelo não impede que, na prática, cada juiz possa adotar um deles, outros modelos de pensamento jurídico ou até mesmo abordagens sincréticas entre os diversos padrões das escolas do Direito. Todavia, o reconhecimento da pluralidade epistêmica advinda a independência funcional de cada juiz (que possibilita maior variabilidade de decisões sobre um mesmo tema) não limita as aparas necessárias às decisões que caminharem para fora dos padrões de racionalidade institucionalmente

⁸⁰⁵ “Nos termos do art. 178 da Constituição da República, as normas e os tratados internacionais limitadores da responsabilidade das transportadoras aéreas de passageiros, especialmente as Convenções de Varsóvia e Montreal, têm prevalência em relação ao Código de Defesa do Consumidor” (BRASIL. Supremo Tribunal Federal (Plenário). **Recurso Extraordinário 636.331**. As Convenções Haia e de Montreal têm prevalência sobre o Código de Defesa do Consumidor e são aplicáveis aos casos de responsabilidade civil em contratos de transporte aéreo internacional. Brasília, DF: Supremo Tribunal Federal, [2017]. Relator: Min. Gilmar Mendes, 25 de maio de 2017. Disponível em: <https://redir.stf.jus.br/paginadorpweb/paginador.jsp?docTP=TP&docID=14028416>. Acesso em: 09 jul. 2022).

⁸⁰⁶ Posteriormente, o Supremo Tribunal Federal criou *distinguishing* em relação à tese do Tema 210 de Repercussão Geral, através da tese de Repercussão Geral de Tema 1.240: “Não se aplicam as Convenções de Varsóvia e Montreal às hipóteses de danos extrapatrimoniais decorrentes de contrato de transporte aéreo internacional” (BRASIL. Supremo Tribunal Federal (Plenário). **Recurso Extraordinário 1.394.401**. Em matéria de danos extrapatrimoniais relacionados ao transporte aéreo internacional, o Código de Defesa do Consumidor tem prevalência sobre as Convenções Varsóvia e de Montreal. Brasília, DF: Supremo Tribunal Federal, [2022]. Relator: Min. Rosa Weber, 15 de dezembro de 2022. Disponível em: <https://portal.stf.jus.br/processos/downloadPeca.asp?id=15356316433&ext=.pdf>. Acesso em: 04 mar. 2023).

⁸⁰⁷ MARQUES, Cláudia Lima. Superação das Antinomias pelo Diálogo das Fontes: O Modelo Brasileiro de Coexistência entre o Código de Defesa do Consumidor e o Código Civil de 2002. **Revista da Escola Superior da Magistratura de Sergipe**, Aracajú, n. 07, 2004.

⁸⁰⁸ GADAMER, Hans-Georg. **Verdade e Método**: traços fundamentais de uma hermenêutica filosófica. 4.^a ed. Petrópolis: Vozes, 2002.

⁸⁰⁹ KELSEN, Hans. **Teoria Pura do Direito**, p. 4-25.

reconhecida, uma vez que os casos podem ser reanalisados nos Tribunais pela via recursal.

Como um padrão de racionalidade a ser incluído nas decisões judiciais auxiliadas por Inteligência Artificial⁸¹⁰ (e suas ontologias⁸¹¹ delimitadas na programação⁸¹²), a teoria da argumentação de Robert Alexy se desenvolve como um processo justificatório a partir de bases racionais, sendo que a mesma se divide em duas fases básicas. A primeira fase é a chamada **justificação interna**⁸¹³, na qual o esforço justificatório passa pela decomposição do conceito trazido em cada norma aplicada, a fim de que de haja a seleção de proposição normativa com o(s) conceito(s) aceito(s). Já a segunda fase é a **justificação externa**, a qual, segundo Alexy, é dividida em “(a) interpretação, (b) argumentação da Ciência do Direito (dogmática), (3) do uso de precedentes, (4) da argumentação prática geral e (5) da argumentação empírica, assim como (6) das chamadas formas especiais de argumentos jurídicos.”⁸¹⁴

Nesse aspecto, embora não seja obrigatória a construção de justificação externa em todas as decisões, haja vista que em decisões mais simples a justificação interna cumpre adequadamente o papel de trazer racionalidade à decisão judicial, em casos de maior densidade argumentativa (e quando exista argumentação de qualquer das partes que direcione o debate) a justificação da decisão judicial deve passar por elementos típicos da ciência do Direito (dogmática).⁸¹⁵ Dessa forma, a ciência do Direito se constitui numa série de **enunciados** doutrinários que se referem à **legislação** e à **aplicação do Direito** (mas que não podem se identificar com a descrição destes), sendo que há, entre si, uma relação de **coerência mútua**, bem

⁸¹⁰ “Ao associar IA com Direito, para além de ferramentas de busca e planejamento estratégico, é muito provável que as soluções que passam por visão computacional, classificações temáticas ou geração de linguagem exerçam, dentro da sua autonomia, algum nível de decisão. A pesquisa e/ou o desenvolvimento, seja na academia, nas empresas ou na administração pública, encontram um objeto de pesquisa. No entanto, no campo do Direito, como ciência social aplicada, haverá níveis variados de dilemas morais” (HARTMANN PEIXOTO, Fabiano; SILVA, Roberta Zumblick Martins da. Op. cit., p. 122).

⁸¹¹ “A statutory ontology contains a taxonomy of normative concepts employed in a statute, their relations to other normative concepts, as well as, their relations to concepts in the regulated domain’s subject matter. In other words, it relates the statutory concepts to the concepts and relations for describing real-world situations” (ASHLEY, Kevin. **Artificial Intelligence and Legal Analytics**, p. 183).

⁸¹² “Uma diferença entre um programador esperto e um programador profissional é que este entende que *clareza é fundamental*. Os profissionais usam seus poderes para o bem, e escrevem códigos que os outros possam entender” (MARTIN, Robert C. **Código Limpo: Habilidades Práticas do Agile Software**. Tradução: Leandro Chu. Rio de Janeiro: Alta Books, 2011, p. 25).

⁸¹³ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 195-203.

⁸¹⁴ Ibid., p. 227.

⁸¹⁵ Alexy utiliza os conceitos de “ciência do Direito” e “dogmática” como sinônimos voltados à classificação dos elementos caracterizadores da teoria do Direito pela doutrina.

como se formam e se discutem tais enunciados dentro de uma dogmática que **funciona institucionalmente** a partir de **conteúdo normativo**.⁸¹⁶

Desse modo, a ciência do Direito se ocupa de enunciados sobre: a) conceitos jurídicos; b) definições de outros conceitos presentes nas normas jurídicas; c) descrições ou caracterizações de estados de coisas; e d) formulação de princípios (enunciados normativos de alto nível de generalidade e que não podem ser aplicados sem se agregar premissas normativas adicionais).⁸¹⁷

O desenvolvimento da ciência do Direito se distingue a) pela linguagem dogmática; b) pelo uso dos enunciados da dogmática; c) pela fundamentação dos enunciados da dogmática (fundamentação dogmática pura⁸¹⁸ e fundamentação dogmática impura⁸¹⁹); e d) pela função da dogmática.

Nesse sentido, verifica-se que a doutrina tem papel central no desenvolvimento do Direito por meio da teoria da argumentação, o que se demonstra através: a) da **função de estabilização** para a fixação dos enunciados dogmáticos, os quais se fazem reproduzíveis; b) da **função de progresso** que possibilita a evolução do Direito no tempo e no espaço; c) da **função de descarga** voltada adoção, nas fundamentações, de enunciados já comprovados e aceitos, pelo menos de maneira provisória; d) da **função técnica** mediante a apresentação simplificada e sistematicamente unificada das normas jurídicas, o que gera a capacidade de transmissão do conhecimento jurídico (função de informação); e) da **função de controle** direcionada à comprovação da consistência dos enunciados dogmáticos (comprovação sistemática em sentido estrito e em sentido amplo); e f) da **função de heurística** que demarca o sistema dogmático como um ponto de partida relativamente seguro para novas observações e relações da ciência do Direito.⁸²⁰

Dessa forma, Alexy destaca, em regra específica de sua teoria da argumentação jurídica, que “se são possíveis argumentos dogmáticos, devem ser usados” (J.12).⁸²¹ Ou seja, a teoria do Direito (com suas desavenças teóricas típicas

⁸¹⁶ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 223-224.

⁸¹⁷ Ibid., p. 224-227.

⁸¹⁸ Quando “o enunciado dogmático a ser fundamentado se segue de outros enunciados dogmáticos, juntamente com enunciados empíricos apenas ou com a adição de normas jurídico-positivas” (Ibid., p. 227).

⁸¹⁹ Casos “em que são necessárias premissas normativas adicionais” e argumentos práticos de tipo geral (Id.).

⁸²⁰ Ibid., p. 232-236.

⁸²¹ Ibid., p. 237.

entre os juristas) é aceita pela teoria da argumentação como parte da racionalidade inerente à decisão judicial, o que se traduz num pluralismo de justificações possíveis dentro do modelo democrático de decisão judicial.

Nesse ponto, é pedagógico o debate acerca da existência de mais de uma teoria do Direito, realizado entre Joseph Raz⁸²², Robert Alexy⁸²³ e Eugenio Bulygin⁸²⁴, o qual demonstra categoricamente a pluralidade de entendimentos a respeito do que é Direito e de quais são seus elementos essenciais. Destaca-se aqui que, em respeito ao pluralismo democrático, reconhece-se a possibilidade de justificação racional das decisões judiciais mediante variados modelos doutrinários de interpretação para um mesmo caso. Todavia, sempre haverá **uma** decisão correta que deverá prevalecer ao final, mediante o modelo que se construir no plano de validação institucional. Garantir o direito de fala do outro⁸²⁵ na decisão judicial auxiliada por Inteligência Artificial é

⁸²² “É possível para qualquer pessoa inventar ou desenvolver um novo conceito. Alguns conceitos que emergem desse modo abrem caminho dentro da cultura geral e usualmente são mais ou menos alterados durante sua trajetória. Mas, em sua maior parte, os conceitos existem independentemente de qualquer de seus usuários. Em geral, aprendemos conceitos em vez de inventá-los ou desenvolvê-los. Tem que ser assim. Dada a riqueza de nossos conceitos, e os limites de nossas capacidades, não é possível para alguém inventar ou modificar mais que uma parte material deles. Dado seu papel na comunicação, inventá-los ou desenvolvê-los frustraria seu propósito. O fato de que, em sua maior parte, os conceitos estão aí, independentemente de nós coletivamente. As condições que fixam a identidade de conceitos particulares são idealizações construídas desde nossas práticas conceituais, quer dizer, desde o uso desses conceitos em geral. Não necessitam refletir nenhuma prática de um indivíduo. Embora seja impossível a existência de um conceito do qual ninguém conheça absolutamente nada é possível que ninguém tenha uma compreensão ou conhecimento completamente corretos de um conceito, ou, de fato, de qualquer conceito, incluindo o conceito de um conceito” (RAZ, Joseph. **Uma Discussão Sobre Teoria do Direito**. Tradução: Sheila Stolz. São Paulo: Marcial Pons, 2013, p. 75).

⁸²³ ALEXY, Robert. Sobre Duas Justaposições: Conceito e Natureza, Direito e Filosofia. Alguns Comentários sobre “Pode Haver Uma Teoria do Direito?” de Joseph Raz. In: RAZ, Joseph. **Uma Discussão Sobre Teoria do Direito**. Tradução: Sheila Stolz. São Paulo: Marcial Pons, 2013, p. 102-108.

⁸²⁴ “Se a lista das propriedades essenciais do direito é indefinida, então cada conceito de direito identificará somente algumas delas, e conseqüentemente, haverá uma pluralidade de conceitos de direito.

Nesse sentido, as diferentes respostas à pergunta ‘o que é direito?’ conduzem a diferentes conceitos de direito. As propriedades necessárias do direito não são as mesmas para Tomás de Aquino, Austin, Holmes, Kelsen, Hart ou Dworkin. A pergunta ‘o que é direito?’ tem recebido numerosas e diferentes respostas ao longo da história e, como Hart corretamente observa, constitui um pedido de uma definição de direito; e o falar da essência ou natureza do direito é só uma variação, obscuramente formulada, da mesma pergunta.

Pelo que, em oposição à Raz, eu sustentaria que, em algum sentido, os diferentes conceitos de direito são produtos da teoria do direito. Eles tratam de elucidar a estrutura da instituição chamada ‘direito’, ainda que de maneiras diferentes” (BULYGIN, Eugenio. Raz e a Teoria do Direito. Comentários Sobre “Pode Haver Uma Teoria do Direito?” de Joseph Raz. In: RAZ, Joseph. **Uma Discussão Sobre Teoria do Direito**. Tradução: Sheila Stolz. São Paulo: Marcial Pons, 2013, p. 117).

⁸²⁵ “É constitutivo para a racionalidade de exteriorização que o falante manifeste uma pretensão de validade criticável em favor de um enunciado ‘p’ que possa ser aceita ou refutada pelo ouvinte” (HABERMAS, Jürgen. **Teoria do Agir Comunicativo**, 1, p. 36).

diferente de aceitar passivamente a decisão, uma vez que os enunciados dogmáticos são refutáveis⁸²⁶, pois “todo enunciado dogmático deve enfrentar uma comprovação sistemática, tanto em sentido estrito como em sentido amplo” (J.11).⁸²⁷

Sendo assim, propostas de implementação de sistemas de Inteligência Artificial judicial que amarrem a pessoa do julgador a uma “camisa de força” de pouca adaptabilidade epistêmico-jurídica tendem à violação da independência dos julgadores.⁸²⁸ Esse risco deve ser seriamente tratado no desenvolvimento de modelos de IA judicial que levem em conta o dever de se garantir a pluralidade de posicionamentos dos magistrados.

O desenvolvimento de sistemas de Inteligência Artificial para auxílio à tomada de decisões judiciais ainda tem um longo caminho, de avanços e cautelas, a trilhar para que se chegue a uma Inteligência Artificial judicial de caráter mais generalista.⁸²⁹ Todavia, o risco à independência⁸³⁰ jurídico-filosófica dos juízes deve ser sopesado desde o presente momento histórico, buscando um modelo pautado na experiência democrática de estruturação do Poder Judiciário.

Dessa maneira, a figura do juiz humano (com sua visão independente acerca da teoria do Direito) deve ser integrada à arquitetura decisória inerente à Inteligência Artificial judicial, haja vista que a figura humana do julgador não deve ser suplantada por um modelo puramente maquinal. A ideia de um “juiz” máquina deve ser rejeitada, já a integração das potencialidades da hibridização do processamento do cérebro humano com a Inteligência Artificial levará à configuração da já referida figura do juiz ciborgue, o que deve ser estimulado.

Nessa linha, um Judiciário integrado por juízes puramente replicadores de posicionamentos pré-programados, sem margem eficiente de adaptabilidade quanto à visão jurídica da ciência do Direito, implicaria numa espécie de democracia de

⁸²⁶ “(J.10) Todo enunciado dogmático, se posto em dúvida, deve ser fundamentado mediante o emprego, pelo menos, de um argumento prático do tipo geral” (ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 231).

⁸²⁷ Id.

⁸²⁸ “(...) só a jurisdição subtraída de influências estranhas pode configurar uma justiça que dê a cada um o que é seu e somente através da garantia de um juiz imparcial o processo pode representar um instrumento não apenas técnico, mas ético também, para a solução dos conflitos interindividuais com justiça” (CINTRA, Antonio Carlos de Araújo; GRINOVER, Ada Pellegrini; DINAMARCO, Cândido Rangel. Op. cit., p. 52).

⁸²⁹ Esse caminho pode, eventualmente, passar pelo uso de sistemas de IA *task agnostic* como o GPT-4 (ChatGPT), por exemplo, com os cuidados que a presente obra busca delinear. Mais sobre esse tema pode ser lido nos subcapítulos 5.7 e 5.8.

⁸³⁰ Art. 35, I, da Lei Complementar 35/1979 e art. 4º ao art. 7º do Código de Ética da Magistratura.

qualidade inferior⁸³¹, uma vez que, à pessoa do juiz, deve ser possibilitada a adaptabilidade na adoção de posicionamento doutrinário, implicando em visão própria do Direito. A aplicação do direito mediante modelo de juiz autômato cumpridor de diretriz centralizada por Tribunal e por órgão de controle (mediante sistema de Inteligência Artificial) traz o risco de que o julgador não se sinta à vontade para julgar de maneira independente.

Soma-se a isso o já comentado fato de que a usabilidade da interface dos programas de computador (que é uma capa para os algoritmos computacionais)⁸³² tem o potencial de levar os juízes a um perigoso estado de acomodação. Tal risco ocorre porque uma interface de fácil utilização nos sistemas de IA judicial pode dificultar a percepção do funcionamento dos algoritmos, o que abre passagem para vieses decisórios trazidos pela programação ou pela base de dados.⁸³³ Os juízes podem ser auxiliados pelos sistemas de Inteligência Artificial judicial⁸³⁴. Contudo, é inconstitucional a transformação dos julgadores em simples marionetes de sistemas de Inteligência Artificial centralizados e movidos apenas através de “um click no computador”⁸³⁵, sem que haja a garantia de um sistema argumentativo democrático de aferição das razões (e dos enunciados dogmáticos) de cada julgamento tomado com o auxílio de IA, conforme indicado no Capítulo 2, ao se delinear o direito fundamental à explicabilidade.

Nesse sentido, ao avançar no tema da indelegabilidade da atividade jurisdicional (vide o Capítulo 1), o art. 17 da Resolução 332/2020 do CNJ é bastante claro ao assegurar a “autonomia dos usuários internos” (conceito no qual se incluem

⁸³¹ “Capturar os árbitros dá ao governo mais que um escudo. Também oferece uma arma poderosa, permitindo que ele imponha a lei de maneira seletiva, punindo oponentes e favorecendo aliados” (LEVITSKY, Steven; ZIBLAT, Daniel. Op. cit., p. 81-85).

⁸³² “*Their transparency is precisely why this class of objects is so appealing to engineers and designers as platforms for computation. These things are already everywhere, hiding in plain sight; nobody bats an eyelash at them. If nothing else, they offer a convenient place to stash the componentry of a computing power that might otherwise read as oppressive*” (GREENFIELD, Adam. Op. cit., p. 19).

⁸³³ DI BELLO, Marcello; VERHEIJ, Bart. Evidence & decision making in the law: theoretical, computational and empirical approaches. **Artificial Intelligence and Law**, Dordrecht, v. 28, 1-5, 2020. DOI: 10.1007/s10506-019-09253-0.

⁸³⁴ PEREIRA, Sebastião Tavares. O *Machine Learning* e o Máximo Apoio Ao Juiz. **Revista Democracia Digital e Governo Eletrônico**, Florianópolis, v. 1, n. 18, p. 2-35, 2019.

⁸³⁵ Por exemplo: “Sessão inédita da 8ª Câmara Cível do Tribunal de Justiça de Minas Gerais (TJMG) julgou, com apenas um click no computador, um total de 280 processos. Em menos de um segundo, todos os processos foram julgados” (MINAS GERAIS. Tribunal de Justiça do Estado de Minas Gerais. **TJMG utiliza inteligência artificial em julgamento virtual**. [2020]. Disponível em: <https://www.tjmg.jus.br/portal-tjmg/noticias/tjmg-utiliza-inteligencia-artificial-em-julgamento-virtual.htm>. Acesso em: 15 set. 2020).

os juízes)⁸³⁶, mediante o uso de modelos de IA que proporcionem o “incremento, e não restrição” (inciso I) e que tragam a possibilidade de revisão da proposta de decisão e dos dados utilizados (inciso II).⁸³⁷

Para além do plano do negacionismo da importância e da aplicabilidade da IA jurídica, tal espécie de posicionamento será enfraquecida com a evolução tecnológica e com fatos que demonstrarão, cada vez mais, a integração necessária entre Inteligência Artificial e decisão judicial.⁸³⁸ Desse modo, por meio dos avanços da tecnologia na interação entre o Direito e Ciência da Computação, haverá a paulatina superação do ceticismo quanto à implementação de modelos de IA judicial mais complexos que possam respeitar a pluralidade de entendimentos (e de argumentações) que os juízes utilizam. Nesse ponto, destaca-se que a Teoria do Direito exerce a sua função epistemológica ao estudar a cooperação Direito-IA de maneira vinculada às tecnologias que já existem. Por outro lado, a Teoria do Direito não pode fechar os olhos para a potencial evolução⁸³⁹ que leve ao desenvolvimento de novos sistemas de IA judicial no futuro. Em analogia, a Física teórica trabalha conceitos e abstrações antes mesmo de existir uma prova científica perceptível de que determinado objeto ou fenômeno exista.⁸⁴⁰ Assim como nem sempre a Ciência é feita em laboratórios, nem sempre o Direito nasce em códigos legais ou institucionais. Esse é um ponto onde a fronteira entre a Teoria do Direito e a Filosofia do Direito é uma linha não demarcada.

⁸³⁶ Segundo o art. 3º, V, da Resolução 332/2020 do CNJ se caracteriza como usuário interno o “membro, servidor ou colaborador do Poder Judiciário que desenvolva ou utilize o sistema inteligente” (CONSELHO NACIONAL DE JUSTIÇA. Resolução nº 332, de 21 de agosto de 2020).

⁸³⁷ Id.

⁸³⁸ *“When these unhuman systems of today are able to make predictions, identify relevant documents, answer questions, and handle emotions (see Chapter 3) at a higher standard than human beings, it is not just reasonable, it is vital that we ask whether, in decades to come, people or systems will be undertaking the wide variety of work that goes on in our courts today”* (SUSSKIND, Richard. **Online Courts and the Future of Justice**, p. 274).

⁸³⁹ “O processo evolucionista da tecnologia melhora a capacidade de modo exponencial. Inovadores procuram melhorar as capacidades por meio de múltiplos. A inovação é multiplicativa, não aditiva. (...) A evolução usa feedback positivo: os métodos mais adequados que resultam de um estágio do progresso evolucionista são usados para criar o estágio seguinte. Conforme descrito no capítulo anterior, cada época da evolução progrediu mais rápido, construindo sobre os produtos do estágio anterior. A evolução trabalha indiretamente: a evolução criou os humanos, os humanos criaram a tecnologia, os humanos estão agora trabalhando com uma tecnologia cada vez mais avançada para criar novas gerações de tecnologia” (KURZWEIL, Ray. Op. cit., p. 61).

⁸⁴⁰ Por exemplo: Peter Riggs, em uma visita à Universidade Princeton, foi tratado com certo desinteresse por Alan Guth quando falou sobre a partícula subatômica, hoje conhecida como *Bóson de Riggs*; Albert Einstein imaginava cavalgar um raio de luz, o que inspirou a Teoria da Relatividade Especial enquanto trabalhava como burocrata no Escritório Federal Suíço Patentes, em Berna; John Wheeler (influenciado pelas pesquisas de diversos outros físicos que o antecederam) cunhou o termo “buraco negro” em 1967, espécie de objeto que teve a sua primeira fotografia obtida no ano de 2019.

Feitas tais considerações, o desenvolvimento de sistemas de Inteligência Artificial judicial deve incluir os juízes no processo de pesquisa e desenvolvimento dos sistemas, assim como, mais do que respeitar o posicionamento de cada magistrado através da possibilidade de abstenção do uso da IA (solução trazida pelo art. 17 e pelo art. 19, Parágrafo único, da Resolução 332/2020 do CNJ⁸⁴¹), deve incluir margem de adaptação aos sistemas de IA judicial, para que possam ser pensados com soluções parametrizadas por cada juiz competente⁸⁴² para os casos que estejam sob sua análise.

A proposta é desafiadora aos desenvolvedores de sistemas, e dois fatores serão decisivos para o enfrentamento do problema: a) a consolidação dos primeiros sistemas de IA judicial (o que inclui a aferição dos resultados da política pública fomentada pelo Conselho Nacional de Justiça); e b) a formação da primeira grande geração de desenvolvedores de IA jurídica (engenheiros de computação jurídica, cientistas de dados jurídicos e outros profissionais nascidos do hibridismo entre Direito e Ciência da Computação⁸⁴³), fenômeno que se encontra em curso a partir da inclusão do estudo de tecnologia nos cursos de graduação e pós-graduação em Direito, os quais serão integrados às equipes de apoio dos magistrados.

A partir dos pontos abordados neste estudo, percebe-se que o art. 21, III, da Resolução 332/2020 do Conselho Nacional de Justiça prevê que não devem existir preconceitos nos estudos, pesquisas, ensino e treinamentos de Inteligência Artificial judicial, sendo vedado “subordinar investigações a sectarismo capaz de direcionar o curso da pesquisa ou seus resultados”.⁸⁴⁴ No entanto, percebe-se ao partir das premissas aqui utilizadas que a referida norma do Conselho Nacional de Justiça não pode ser interpretada de maneira a considerar sectarismo a busca de implementar sistemas de Inteligência Artificial judicial voltados ao respeito ao pensamento jurídico de cada juiz dentro do âmbito de sua independência decisória.

Demonstra-se o risco de desrespeito à independência judicial por meio da captura dos Juízes por sistemas de Inteligência Artificial desenvolvidos a partir de

⁸⁴¹ CONSELHO NACIONAL DE JUSTIÇA. Resolução nº 332, de 21 de agosto de 2020.

⁸⁴² “Os desenvolvedores devem se sensibilizar com os indícios que revelam a presença de conceitos implícitos e, às vezes, é preciso procurá-los de forma proativa. A maioria dessas descobertas vem de escutar a linguagem da equipe, examinar minuciosamente as estranhezas no design e as aparentes contradições nas afirmações dos especialistas, esmiuçar a literatura sobre o domínio e fazer várias e várias experimentações” (EVANS, Eric. **Domain-Driven Design**, p. 198).

⁸⁴³ SUSSKIND, Richard. **Tomorrow's Lawyers**, p. 135-143.

⁸⁴⁴ CONSELHO NACIONAL DE JUSTIÇA. Resolução nº 332, de 21 de agosto de 2020.

visão extremamente fechada ou muito particular da Teoria do Direito, como, por exemplo, a busca por se vincular as decisões judiciais às concepções originalistas⁸⁴⁵ ou exegéticas de aplicação de texto legal sem a observância do contexto, tornando ausente qualquer margem de interpretação ao julgador.

Dessa maneira, a legitimidade da utilização de modelos de Inteligência Artificial judicial se desenvolve através de um processo comunicacional cooperativo, um duplo canal de correção, voltado, de um lado, à diminuição de ruídos e vieses e, de outro lado, ao incentivo às decisões judiciais corretas. Esse duplo canal de correção é alcançado pela pretensão de correção⁸⁴⁶ em dois níveis. O primeiro nível advém dos sistemas de Inteligência Artificial que atuam como meios comunicacionais de proposições normativas⁸⁴⁷, previamente programadas ou inferidas a partir dos dados, de maneira a enriquecer o processo argumentativo humano. O segundo nível consiste na pretensão de correção buscada por um julgador racional e independente na aplicação do Direito.

Assim, há uma cooperação dialético-argumentativa na qual os enunciados doutrinários advindos da ciência do Direito (dogmática) integram o processo argumentativo (justificação externa) de maneira legítima, motivo pelo qual devem ser incluídos (ou levados em conta) no desenvolvimento de sistemas de Inteligência Artificial judicial, sendo que concomitantemente os sistemas de IA judicial podem sugerir novas abordagens ou questionar os argumentos trazidos pelo Juiz humano. Logo, pode haver uma “dança” argumentativa bem coordenada nessa interação Juiz-IA, na qual, sucessivamente, humano e máquina buscam demonstrar as melhores razões para uma decisão, mediante recíproca correção, até que o julgador humano tome a decisão final adotando os argumentos racionais que lhe foram disponibilizados pela interação humano-máquina. Essa é a atuação da pluralidade epistêmica através do duplo canal de correção. O juiz ciborgue assim se lapida, com computação consciente e razão computada, na busca da decisão correta.

A pluralidade de visões acerca da Teoria do Direito é garantida aos julgadores por meio de seu dever de independência, o que se traduz num direito fundamental à

⁸⁴⁵ Por exemplo, na época da promulgação da Constituição (1988) e quando foi aprovado o Código Civil (2002) não se cogitava de propriedade digital, o que poderia levar um sistema de automatização de julgamentos com viés originalista a desconsiderar totalmente tal possibilidade, sem estímulo ao processo de individualização da norma jurídica no novo contexto vivido atualmente.

⁸⁴⁶ ALEXY, Robert. **Conceito e Validade do Direito**, p. 154-155.

⁸⁴⁷ BULYGIN, Eugenio. **Lógica Deontica, Normas y Propositiones Normativas**, p. 60-62.

existência de juízes independentes (art. 5º, XXXV, da Constituição Federal) que tenham a possibilidade de, quando necessário, debaterem a respeito das suas diferentes concepções de Direito no sentido de influenciar o resultado das decisões⁸⁴⁸ (baseadas em diversas angulações da Teoria do Direito), o que ocorre por meio de um processo argumentativo garantido pelo duplo canal de correção comunicacional Juiz-IA.

A pesquisa e o desenvolvimento dos sistemas de Inteligência Artificial voltados à decisão judicial têm desafios inerentes à complexidade do tema, o que não impede a adoção ética de modelos de IA judicial menos complexos, que respeitem a independência judicial, passando gradualmente para modelos de IA jurídica mais complexos (com a adoção de elementos de Teoria do Direito em sua programação e com representação adequada na base de dados) que possibilitem a adaptabilidade dos sistemas computacionais inteligentes ao pensamento jurídico-epistêmico de cada juiz, garantindo, assim, o *due process of law* da era da IA judicial.

Nesse caminho de desenvolvimento, o duplo canal de correção advindo da interação comunicacional entre julgador e sistema de IA judicial se conecta ao aspecto democrático, pautado no pluralismo, que leva ao enriquecimento epistêmico de uma sociedade plural. A partir disso, conforme referido ao se tratar do direito fundamental à explicabilidade (Capítulo 2), a concepção da decisão judicial auxiliada por IA também pode trazer os argumentos (do Juiz humano ou da Inteligência Artificial) que foram superados no processo de justificação (de maneira análoga àquilo que ocorre em relação às divergências em julgamentos colegiados). Uma vez vencidos, ao serem devidamente registrados e publicizados, os argumentos têm o potencial de serem revisitados em novos momentos históricos⁸⁴⁹ e ainda podem ser utilizados em casos futuros por outros julgadores e por atores da sociedade.⁸⁵⁰

⁸⁴⁸ DWORKIN, Ronald. **O Império do Direito**. 3ª ed., p. 20-29.

⁸⁴⁹ “That is a comforting story-except for the fact that there are always dissenters, and we never know which is going to be vindicated and which repudiated by history. Once we realize, as we must, that some dissenters were wrong at the moment they dissented and wrong thereafter, the story about dissent, and about how history decides who is the winner and who the loser, becomes much more complicated” (TUSHNET, Mark. **I Dissent: Great Opposing Opinions in Landmark Supreme Court Cases**. Boston: Beacon Press, 2008, posição 25-28. E-book Kindle).

⁸⁵⁰ “So, dissents can matter, but in quite an indirect way: a dissent sent might be picked up by a social movement because the dissent sent expresses something the movement already has in its constitutional vision; the social movement's constitutional vision might affect a political party and its candidates; successful candidates might nominate judges and justices because of their constitutional visions; and these new justices might conclude that the dissent-by now, perhaps years in the past-provides a better account of our Constitution than the majority opinion does. All of the qualifications are important here” (Ibid., posição 115-119).

3.4 Pré-Compromissos, Heurísticas e Inteligência Artificial Judicial

3.4.1 A Caracterização dos Consensos como Heurísticas para a IA Judicial

Conforme já exposto, diversos Tribunais brasileiros vêm aprofundando pesquisas sobre a relação entre Inteligência Artificial e direito, no intuito de criar sistemas que possam, inclusive, propor modelos de decisões aos juízes.

Assim, para aprofundar o debate sobre a temática, adotou-se o modelo de argumentação jurídica de Robert Alexy, focando, especificamente, no que o autor denomina como fase da justificação interna. A escolha do paradigma da justificação interna passa pela necessidade de investigar a interação entre argumentação jurídica e Inteligência Artificial desde seus aspectos mais basilares, a fim de estabelecer uma linha demarcatória entre aquilo que é plausível à racionalidade do Direito e modelos de simples proposição matemática de soluções jurídicas por meio de computação preditiva.

Dessa forma, há um esforço investigativo com enfoque jurídico sobre a tecnologia da Inteligência Artificial (representação do conhecimento, *machine learning* e redes neurais, por exemplo) e sua aplicabilidade para a estruturação da argumentação jurídica.

Em reforço a essa perspectiva, seria inadequado negar os milhões de anos⁸⁵¹ de evolução da biologia do pensamento humano⁸⁵² mediante teoria que encobrisse tal característica, pois a estrutura do pensamento humano possui elementos próprios de funcionamento (incluindo-se a “síntese”⁸⁵³ do mundo perceptível), os quais são pressupostos à própria existência da decisão judicial enquanto uma atividade humana.

⁸⁵¹ Cerca de 2 milhões de anos, considerando-se a evolução do cérebro a partir do *Homo habilis* até o *Homo sapiens*.

⁸⁵² HARARI, Yuval Noah. **Sapiens**, posição 148-152.

⁸⁵³ “É a síntese em geral, como proximamente veremos, a simples obra da imaginação, quer dizer, uma função cega, ainda que indispensável, da alma, sem a qual não teríamos conhecimento de nada, função de que raras vezes temos consciência. Mas é uma função que pertence ao entendimento, e que é a única que nos procura o conhecimento propriamente dito, o reduzir esta síntese a conceitos.

A síntese pura, representada geralmente, nos dá o conceito intelectual. Mas entendo por síntese pura, a que se funda em um princípio da unidade sintética ‘a priori’. Assim nossa numeração (o que se nota melhor ainda nos números elevados) é uma síntese segundo conceitos, porque tem lugar segundo um princípio comum de unidade (p. ex.: o decimal). Sob esse conceito é necessária a unidade na síntese da diversidade. Podem submeter-se, mediante a análise, diferentes representações a um só conceito, assunto de que se ocupa a Lógica geral. A Lógica transcendental, pelo contrário, ensina a submissão aos conceitos, não das representações, mas da síntese pura das representações” (KANT, Immanuel. **A Crítica da Razão Pura**, posição 1430-1437).

Os estudos acerca dos processos de decisão não são uma exclusividade do Direito, pois outros ramos do conhecimento se interessam profundamente pelo tema, como a Psicologia, a Filosofia, a Administração e a Ciência da Computação, por exemplo.⁸⁵⁴

Nesse sentido, o campo de estudo é vasto e qualquer pretensão de revisitar todos os temas-chave seria ingênua. Assim, como lembra Nassim Taleb⁸⁵⁵, deve-se desconfiar de teorias com pretensão de completude, haja vista a opacidade da realidade frente à percepção do intelecto humano. Dessa forma, falar sobre pré-compromissos e heurísticas como elementos basilares para o desenvolvimento de sistemas de IA judicial é flertar com a incompletude das soluções pensadas no tecido do processo político⁸⁵⁶, com o limite do conhecimento e com aquilo que já se conhece para não tropeçar nos obstáculos do desconhecido⁸⁵⁷, bem como a dificuldade de computação do mundo jurídico (e de sua natureza entrópica, conforme descrito no Capítulo 1).

Ainda que o direito tenha que empregar um importante esforço para se adaptar às novas realidades e aos novos ares, deve-se ter em mente que na ponderação jurídica primordial entre justiça e segurança jurídica⁸⁵⁸, esta última também tem seu espaço, seu momento e sua função no regime democrático. Para além de ser uma norma inerente ao direito fundamental à segurança jurídica, conforme se extrai do art. 5º, *caput*, da Constituição Federal e de outras normas do ordenamento jurídico brasileiro, a estabilização das relações jurídicas é um valor⁸⁵⁹ que deve ser perseguido na democracia.⁸⁶⁰

Nessa linha, serão analisados os principais tipos de heurísticas existentes⁸⁶¹ e a influência dos modelos heurísticos na formação de pré-compromissos decisórios⁸⁶²

⁸⁵⁴ Neste aspecto, não é objetivo desta obra explicar a questão complexa do pensamento humano em seus mais variados aspectos. O enfoque do estudo do pensamento humano está em como determinados elementos da teoria da racionalidade limitada (*bounded rationality*) podem ajudar a compreender e a aprimorar a interpretação do Direito.

⁸⁵⁵ TALEB, Nassim N. Op. cit., p. 37.

⁸⁵⁶ WALDRON, Jeremy. **Law and Disagreement**. New York: Oxford, 2004, p. 256.

⁸⁵⁷ THALER, Richard H.; SUNSTEIN, Cass R. Op. cit., posição 466-467.

⁸⁵⁸ ALEXY, Robert. **Teoria de Los Derechos Fundamentales**, p. 138.

⁸⁵⁹ Ibid., p. 143.

⁸⁶⁰ HABERMAS, Jürgen. **Direito e Democracia**, p. 316-317.

⁸⁶¹ THALER, Richard H.; SUNSTEIN, Cass R. Op. cit., p. 423-426.

⁸⁶² WALDRON, Jeremy. Op. cit., p. 257.

voltados à composição de argumentação⁸⁶³ apta a influenciar o processo de decisão auxiliado por Inteligência Artificial judicial.

Em tal aspecto, Waldron adota a ideia do pré-compromisso como a forma de vinculação à segurança (inclusive contra maiorias legislativas eventuais) e à previsibilidade de aplicação do Direito⁸⁶⁴, num balanço delicado entre democracia e soberania popular.⁸⁶⁵ Em sua teoria, pode-se perceber a distinção entre o pré-compromisso individual (cujo grau de consenso é mais tênue) e o pré-compromisso democrático, de caráter vinculativo mais denso.⁸⁶⁶ Todavia, vale destacar que, a depender de cada desenho institucional do respectivo regime constitucional, os pré-compromissos variam na forma de vinculação e nos elementos necessários à revisão dos mesmos.⁸⁶⁷

Um pré-compromisso individual⁸⁶⁸ pode ser um mecanismo causal (espécie de mecanismo à prova de falhas) ou pode se constituir por uma limitação prévia mediante a atribuição de uma decisão ao julgamento de outra pessoa que não seja a afetada diretamente pelo resultado. Nesse sentido, em caso de pré-compromisso individual, o agente é considerado bastante certo (em seus momentos de lucidez) sobre as ações que deseja evitar, pelo fato dessas ações serem indesejáveis em si mesmas no contexto do tecido social. Nessa linha, o agente tem percepção a respeito das ações que ele possa ser tentado a executar no futuro (por exemplo: o fumante que conhece os males do cigarro; e o motorista que entrega as chaves de seu veículo a um amigo, pois sabe que é perigoso dirigir sob efeito de álcool).⁸⁶⁹ Por outro lado, os pré-compromissos democráticos representam uma problemática que demanda uma análise mais detalhada.

Assim, questão reside na ponderação entre a vontade do presente (e seu poder de vinculação do futuro mediante a predição de situações de risco ou indesejadas e o

⁸⁶³ HABERMAS, Jürgen. **Direito e Democracia**, p. 241.

⁸⁶⁴ WALDRON, Jeremy. Op. cit., p. 257-258.

⁸⁶⁵ Ibid., p. 256-257.

⁸⁶⁶ Ibid., p. 271.

⁸⁶⁷ DIXON, Rosalind. Argumento Central a Favor da Forma Fraca do Controle de Constitucionalidade. **Revista Direitos Fundamentais & Democracia**, Curitiba, v. 24, n. 2, p. 5-55, mai./ago. 2019. DOI: 10.25192/issn.1982-0496.rdfd.v24i21647.

⁸⁶⁸ WALDRON, Jeremy. Op. cit., p. 260-263.

⁸⁶⁹ Ibid., p. 259.

planejamento de ações) e a vontade do futuro, a qual poderá buscar se expor à uma situação de risco ou previamente indesejada.⁸⁷⁰

Por sua natureza, o consenso⁸⁷¹ é intergeracional na medida em que o “eu do presente” adota medidas de contenção de eventuais desvios e iniquidades a serem potencialmente praticados pelo “eu do futuro”. Por outro lado, o “eu do futuro” (quando já estiver consciente que é o novo “eu do presente”) pode não mais concordar com os pré-compromissos antes assumidos no nível democrático (constitucional, legal ou em relação aos precedentes judiciais), o que implica na potencial superação do, neste momento, “eu do passado”. Do mesmo modo, um sistema de IA judicial pode atuar como um mecanismo de alerta inerente ao duplo canal de comunicação Juiz-IA, afim de que, na busca da decisão correta, o julgador tome ciência dos pré-compromissos e possa considerá-los na construção argumentativa das suas decisões.

Ademais, Habermas⁸⁷² lembra que os consensos possuem um grau de provisoriedade, pois a coação do melhor argumento pode mostrar ser necessária a repactuação do consenso constitucional em algum momento no futuro.⁸⁷³ Tal perspectiva de Habermas se restringe aos aspectos procedimentais do uso público da razão e desenvolve “o sistema dos direitos a partir da ideia de sua institucionalização legal”⁸⁷⁴, o que passa pelas “condições para discursos e negociações racionais”.⁸⁷⁵ Nesse sentido, sempre podem existir adaptações ou alterações nos sistemas de Inteligência Artificial judicial, de modo a melhor representar os consensos democráticos e as suas influências para a formação de decisões judiciais auxiliadas por sistemas de Inteligência Artificial.

Sobre os consensos democráticos, percebe-se que tanto a leitura de Waldron quanto à abordagem de Habermas são posições mais sensíveis à realidade de conformação político-jurídica dos estados democráticos, indicando modelos que refletem a democracia em seu dinamismo de profusão de ideias.

Essa abordagem possibilita uma viagem no espaço-tempo jurídico para criar freios de contenção aos abusos da vontade futura contaminada pela visão tópica (e potencialmente tóxica) do momento em que se desenvolverá a decisão. Esse

⁸⁷⁰ CONSANI, Cristina Foroni. A Crítica de Jeremy Waldron ao Constitucionalismo Contemporâneo. **Revista da Faculdade de Direito – UFPR**, Curitiba, v. 59, n. 2, p. 143-173, 2014.

⁸⁷¹ WALDRON, Jeremy. Op. cit., p. 270-271.

⁸⁷² HABERMAS, Jürgen. **Teoria do Agir Comunicativo**, 1, p. 61.

⁸⁷³ HABERMAS, Jürgen. **A Inclusão do Outro**, p. 73.

⁸⁷⁴ Ibid., p. 87.

⁸⁷⁵ Ibid., p. 88.

pensamento tem como pressuposto a possibilidade de pré-compromissos desejáveis e necessários para o funcionamento da democracia. Contudo, os pré-compromissos podem sofrer desgaste temporal e passarem por pressão político-jurídica para sua ressignificação, ou mesmo para sua extinção. Aliás, essa entropia é natural, pois tudo que é organizado, desde a matéria⁸⁷⁶ até a informação⁸⁷⁷, tende a se desorganizar com o passar do tempo, esse imperativo não discrimina nada e alcança inclusive o Direito. Todo consenso tenderá, em algum momento, a uma nova rodada de repactuação causada pelo dissenso que surge na vida democrática.

Dessa maneira, no futuro, a sociedade pode ter posicionamentos democraticamente fracos, oportunidade na qual, no âmbito jurídico, os pré-compromissos podem ser ativados para proteger a sociedade em seu momento de fragilidade, a fim de que, como no exemplo de Ulisses tentado pelo canto das sereias, a democracia resista à sua corrosão interna em relação às instituições básicas de funcionamento.⁸⁷⁸ Essa resiliência também deve alcançar o Direito, fruto do regime democrático, e não inibe rearranjos para a recomposições de interesses constitucionais de minorias⁸⁷⁹, por exemplo. Assim, os pré-compromissos democráticos definem a maquinaria institucional de funcionamento do Estado e estabelecem um rito constitucional mais dificultoso ao processo de emenda, o que atua sobre os problemas de paixão, inconsistência e eficiência na aplicação do Direito.⁸⁸⁰

⁸⁷⁶ CHAKRABARTI, C.; CHAKRABARTY, Indranil. Boltzmann Entropy: Probability and Information. **Romanian Journal of Physics**, Bucharest, v. 52, n. 5–7, p. 559–564, 2007.

⁸⁷⁷ SHANNON, Claude E. A Mathematical Theory of Communication.

⁸⁷⁸ ELSTER, Jon. Ulysses and the sirens: A theory of imperfect rationality.

⁸⁷⁹ *“Even considered as a horizontal assemblage of individuals or groups society may have some ability to bind itself, in a sense that is partly literal and partly metaphorical. If the precommitment decision is unanimous, it has better claims to be an act of self-binding than if it is a case of a majority imposing its constitutional views on a minority. And even in the latter case, the minority might prefer the majority's constitution to a purely statutory system. If the problems to which the constitution is offered as a solution can be expected to persist indefinitely, the framers can say with some justification that they are acting on behalf of a temporally extended “self” that also includes future generations. Yet these conditions are far from always met. Frequently, constitutions are imposed on minorities and on future generations in the interest of a majority of the founding generation. Also, many apparent acts of selfbinding turn out, on closer inspection, to be motivated by partisan motives”* (ELSTER, Jon. **Ulysses Unbound: Studies in Rationality, Precommitment, and Constraints**. Cambridge: Cambridge University Press, 2000, p. 168. E-book Kindle).

⁸⁸⁰ *“Constitutional precommitment, therefore, works at two levels. At a first level, the constitution may design the ordinary machinery of government so as to counteract passion, overcome time-inconsistency, and promote efficiency. At a higher level, the machinery of amending the constitution itself may be designed to be slow and cumbersome. These higher-level constraints have two effects. On the one hand, they act directly on the problems of passion, time-inconsistency, and efficiency. On the other hand, they underwrite and stabilize first-level mechanisms that, in turn, act on the same problems”* (Ibid., p. 117-118).

Nessa linha, há uma distinção entre democracia e soberania popular, uma vez que as decisões que determinam a forma da constituição institucional básica de uma sociedade são tecidas no contexto da vida política comum, decisões essas que, na medida em que ocorrem, podem ser indistinguíveis do cotidiano da tomada de decisões políticas próprias das instituições democráticas.⁸⁸¹ Portanto, uma atribuição da soberania popular a um sistema político é uma questão de julgamento. Exige-se o descobrimento de quais decisões contam como constitutivas do sistema político em questão, bem como é necessário que se arrisque certa hipótese explicativa para demonstrar o significado constitucional das decisões, de maneira a contar com a hipótese no sentido de que elas permaneceram ou se estabeleceram como práticas constitucionais porque eram aceitáveis como tal para a maioria dos membros comuns da sociedade.⁸⁸²

Nesse ponto, partindo da concepção dos pré-compromissos estabelecidos como proteção dos julgamentos às paixões humanas e à inconsistência⁸⁸³, restrição advinda de nível de proteção que deve ser buscado na IA judicial a partir de uma ideia de racionalidade limitada, é necessário perceber que a existência de restrições constitucionais escritas e o poder dos tribunais para interpretá-las e aplicá-las são mecanismos de limitação que as pessoas deliberadamente e por boas razões optaram por impor.⁸⁸⁴ Isso se verifica desde de uma “pretensão de fundamentalidade”⁸⁸⁵, que se mostra um processo de arquitetura de escolhas para a melhoria da qualidade das decisões dentro do necessário desenho institucional, pois um “arquiteto de escolhas tem a responsabilidade de organizar o contexto no qual as pessoas tomam decisões”.⁸⁸⁶

Dessa forma, a inteligência desse tipo de autocontrole consiste no aspecto de que a legislação majoritária pode vir a ser injusta, mediante a discriminação de minorias, que são impotentes politicamente para resistir, pela maioria. Assim, os pré-

⁸⁸¹ WALDRON, Jeremy. Op. cit., p. 256.

⁸⁸² Ibid., p. 257.

⁸⁸³ ELSTER, Jon. **Ulysses Unbound**, p. 7-45.

⁸⁸⁴ WALDRON, Jeremy. Op. cit., p. 257.

⁸⁸⁵ “Quem afirma algo, não só quer expressar que acredita em algo, mas também que pretende que o que ele diz é fundamentável, que é verdadeiro, que é correto. Isso vale tanto para proposições normativas como para as não normativas” (ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 171.

⁸⁸⁶ THALER, Richard H.; SUNSTEIN, Cass R. Op. cit., posição 89-90.

compromissos se constituem em heurísticas⁸⁸⁷ que restringem a possibilidade de pura e simples revisão da interpretação pelo julgador.

Com foco nessa concepção de racionalidade limitada, essa heurística de julgamento, caracterizada como um ponto necessário para o sistema democrático, reverte-se em respeito aos direitos individuais, definindo-se como incompatível com uma imagem puramente predatória de maiorias legislativas eventuais. Dessa forma, se as restrições constitucionais e os mecanismos de *judicial review* forem estabelecidos por decisão popular, os mesmos podem ser vistos como precauções tomadas pelos agentes decisórios contra suas próprias imperfeições (por exemplo: Ulisses que se amarra ao mastro do navio; ou o motorista que esconde a chave quando irá experimentar bebida alcoólica). Tais arranjos relacionam-se com a ideia de autogoverno, não caracterizando simples derrogações da liberdade individual.

Portanto, as heurísticas levam o sistema jurídico a “esquecer” o calor do debate majoritário e a demarcar o terreno, de maneira a garantir a adoção de comportamento decisório baseado em arquitetura inteligente⁸⁸⁸ que estimula um ambiente de decisões democráticas.

A partir dessa abordagem, embora não haja uma garantia de que no momento original do estabelecimento dos pré-compromissos exista uma racionalidade superior em relação ao momento de paixão que emerge da tensão posterior, essa superação da racionalidade anterior (pré-compromisso) ocorre pelas vias político-institucionais⁸⁸⁹, nas quais se insere a legitimidade do controle de constitucionalidade exercido pelos juízes⁸⁹⁰ (ainda que este tenha suas limitações). Logo, descabe aos sistemas de Inteligência Artificial judicial realizar proposições que ultrapassem os limites jurídicos institucionalmente estabelecidos, ainda que os julgadores humanos tenham a possibilidade de atuar nesse sentido, em decorrência da já referida pluralidade epistêmica inerente à sua independência (vide o Capítulo 2).

⁸⁸⁷ KAHNEMAN, Daniel. Op. cit., posição 2037.

⁸⁸⁸ THALER, Richard H.; SUNSTEIN, Cass R. Op. cit., posição 730-738.

⁸⁸⁹ “Uma vez que esse cenário é reconhecido, pode-se observar o anverso da imagem de um suposto acordo ou consenso definitivo, o qual procura dar continuidade e (artificialmente) constituir a realidade social. Fazem-se possíveis a revisão, o abandono de práticas anteriores, a acolhida de novas situações em que renovados compromissos podem ser realizados. Por isso, os desentendimentos, os desacordos não são observados como um ‘canto de sereia’, como uma forma patológica, mas como constitutivos do *ethos* da política, de uma comunidade preocupada com a democracia, com a possibilidade da autodeterminação” (LORENZETTO, Bruno Meneses. **Os Caminhos do Constitucionalismo para a Democracia**, p. 244-245).

⁸⁹⁰ O que denota a existência de interpretação constitucional evolutiva inerente à mutação constitucional, por exemplo.

Noutro enfoque, deve-se perceber que a arquitetura de escolhas não necessariamente estimula decisões democráticas, pois para tanto o processo de escolha da arquitetura em si deve ter como pressuposto à formação de pré-compromissos através de consensos democráticos. Sendo assim, observados os critérios democráticos de formação dos pré-compromissos que se manifestam através do Direito (normas constitucionais, leis, precedentes, etc.), é possível utilizá-los como elementos de racionalidade limitada (*bounded rationality*) a fim de nortear as proposições normativas através de sistemas de Inteligência Artificial judicial.

Nesse caminho, o ato de pré-compromisso pode ser autônomo, contudo, a autonomia do agente A que estabelece o pré-compromisso (na época t1), pode ser limitada pelo julgamento do agente B, que aplica o pré-compromisso na época futura t2.⁸⁹¹ Assim, a dimensão interpretativa pode atuar dentro do espaço que encontra na textura da linguagem⁸⁹² e na multiplicidade de significados⁸⁹³ que o próprio pré-compromisso limita ao intérprete no julgamento.

No exemplo do motorista, o agente sabe que, após ingerir álcool, seu julgamento sobre a sua capacidade de dirigir com segurança será seriamente prejudicada. Esse caso se enquadra na categoria de patologia decisória ou *akrasia*.⁸⁹⁴ Esta categoria é inerente às restrições constitucionais e ao controle do aspecto político no regime democrático, pois, guardadas as proporções limitadas ao exemplo, uma vontade da sociedade futura pode representar o rompimento com os consensos sobre as formas de vida e de organização estabelecidas como importantes para a própria existência da sociedade, o que, desde já, demonstra que laços de coesão jurídico-social não podem ser revistos sem prévio debate democrático e institucional. Tais revisões, a depender do contexto histórico, podem ser relativamente traumáticas (conforme for a profundidade da ruptura nas bases do consenso constitucional⁸⁹⁵) ou simplesmente decorrer da maior criatividade para o desenvolvimento da vida social ao longo das gerações.⁸⁹⁶

⁸⁹¹ WALDRON, Jeremy. Op. cit., p. 262.

⁸⁹² HART, Herbert A. L. Op. cit., p. 164.

⁸⁹³ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 199.

⁸⁹⁴ WALDRON, Jeremy. Op. cit., p. 266.

⁸⁹⁵ RAWLS, John. **O Liberalismo Político**, p. 205-211.

⁸⁹⁶ “When an artist chooses to be constrained, we must assume it is because he believes he will benefit artistically from having a smaller choice set. These are cases of standard or intentional precommitment as defined in Chapter Li. When the constraints are imposed from the outside, he may or may not benefit. If he does, we are dealing with incidental constraint as defined in Li. Sometimes, an incidental constraint may turn into an essential one, if the artist chooses to abide by the constraint even when it is no longer

A figura da *akrasia*, conforme o pensamento de Waldron, representa a crise do consenso intertemporal, pois a *akrasia* desafia as heurísticas estabelecidas de maneira a implicar em rodada de decisão desorientada, movendo a decisão para longe do padrão previamente gravado no sistema jurídico, motivo pelo qual é vício decisório que deve ser combatido. Desse modo, para além da destruição das heurísticas decisórias, o esquecimento da história invariavelmente tem o potencial de superar pré-compromissos estabelecidos no nível jurídico-interpretativo, desafiando o próprio regime democrático em sua demanda por previsibilidade e coerência interna. Nesse ponto, os sistemas de Inteligência Artificial judicial podem ser inseridos como um elemento de coesão interpretativa, pois ainda que não possam ser enquadrados como um mecanismo à prova de falhas (vide o exemplo do motorista, referido acima), exigirão do julgador a superação argumentativa humana dos argumentos trazidos por meio de IA.

Por outro lado, o cerne da teoria dos precedentes⁸⁹⁷, que não é o tema principal desta obra, aponta que a superação de posicionamento anterior sem a existência de um debate nos planos político-jurídico (mediante *overruling*) e intertemporal (por meio da historicidade) implica em decisão de pior qualidade, pois o que se tem é a formação de um novo consenso sem que haja a retomada do consenso anterior no âmbito institucional, o que denota ausência de integridade.⁸⁹⁸ Nesse ponto, cabe ressaltar que o funcionamento da democracia se dá por meio de instituições⁸⁹⁹, e que mesmo a criação argumentativa de nova interpretação para embasar decisão judicial deve ser construída mediante argumentação jurídica institucionalizada⁹⁰⁰, institucionalização

mandatory. Film directors may abstain from using sound and color even when these technologies are available.

*Constraints must leave room for choice. For there to be something for the artist to create, the work of art must not be like a crossword puzzle in which there is one and (ideally) only one arrangement of letters that satisfies all the constraints. The creation of a work of art can in fact be envisaged as a two-step process: choice of constraints followed by choice within constraints. The interplay and back-and-forth between these two choices is a central feature of artistic creation, in the sense that choices made within the constraints may induce the artist to go back and revise the constraints themselves” (ELSTER, Jon. **Ulysses Unbound**, p. 176).*

⁸⁹⁷ BARBOZA, Estefânia Maria de Queiroz. Op. cit., p. 180-194.

⁸⁹⁸ LORENZETTO, Bruno Meneses; CLÈVE, Clèmerson Merlin. Interpretação Constitucional: entre dinâmica e integridade. **Seqüência**, Florianópolis, n. 72, p. 67-92, abr. 2016. DOI: <http://dx.doi.org/10.5007/2177-7055.2016v37n72p67>.

⁸⁹⁹ SÁNCHEZ-CUENCA, Ignacio. Power, Rules, and Compliance. In: MARAVALL, José María; PRZEWORSKI, Adam (eds.). **Democracy and the Rule of Law**. Cambridge: Cambridge University Press, 2003, p. 80-81.

⁹⁰⁰ MACCORMICK, Neil. **Argumentação Jurídica e Teoria do Direito**, p.13-15.

esta que deve ser retratada, o mais fielmente possível, nos sistemas de Inteligência Artificial judicial.

Nessas circunstâncias, embora um pré-compromisso constitucional não seja garantia de imutabilidade da interpretação jurídica estabelecida, o consenso político-jurídico estabelecido é um marco que somente pode ser revisto por meios institucionais democráticos. Desse modo, num plano ideal, a IA judicial deve prover ao julgador recursos computacionais voltados à estabilidade e à adaptabilidade jurídicas, concomitantemente. A partir desse dilema inerente à crise do pré-compromisso, Waldron utiliza o exemplo da sua personagem Bridget, onde o “eu do presente” (que mudou de religião) proibiu um comportamento ao “eu do futuro” (acessar a biblioteca teológica) e estabeleceu um pré-compromisso com um órgão de controle (entrega a chave da biblioteca para amigos). De maneira superveniente, o “eu do futuro” (agora no presente) pretende realizar a quebra do pré-compromisso com o “eu do passado” (mediante a devolução das chaves). Nesse exemplo, o órgão de controle, representado pelos amigos de Bridget, deve devolver as chaves? Ou, no contexto do estudo aqui focado, a IA deve lembrar o julgador que a “chave” para essa biblioteca proibida ainda existe? Dessa forma, o exemplo é bastante realista ao reconhecer que há um limite aos pré-compromissos, sendo que no espaço democrático, no limite da resiliência do consenso, o pré-compromisso deve ser revisto, abrindo-se espaço para os desacordos na interpretação jurídica e para a política no ambiente democrático.

Logo, no que tange às escolhas político-democráticas, as restrições constitucionais caracterizam pré-compromissos menos irreais e mais aceitáveis⁹⁰¹, o que é possível, em certos casos na realidade brasileira (observadas as cláusulas pétreas do art. 60, §4º, da Constituição Federal), pelas restrições referidas através de emendas constitucionais (ainda que isso seja politicamente dificultoso). Isso ocorre porque a objeção democrática seria mais forte se a oportunidade de emenda à constituição não existisse.⁹⁰² A partir dessa abordagem, por mais aberto e pluralista que deva ser o desenvolvimento da IA voltada ao auxílio ao julgamento, não devem existir sugestões de decisões baseadas em interpretações jurídicas que neguem validade às cláusulas pétreas do art. 60, §4º, da Constituição Federal. Tal diretriz foi

⁹⁰¹ WALDRON, Jeremy. Op. cit., p. 262.

⁹⁰² Ibid., p. 275.

corretamente incorporada pela Res. 332/2020 do Conselho Nacional de Justiça, na busca de desenvolvimento, implantação e uso de Inteligência Artificial judicial mediante “compatibilidade com os Direitos Fundamentais” (art. 4º) e garantia à “segurança jurídica” (art. 5º).⁹⁰³ Nesse sentido, as cláusulas pétreas estão mais ligadas ao exemplo de pré-compromisso individual (mecanismo à prova de falhas), um mastro para que Ulisses fique firme ao ouvir o sedutor canto das sereias.⁹⁰⁴

Pode-se observar que as restrições constitucionais são externas às instituições democráticas criadas para agir em nome do povo, mesmo que as restrições sejam inerentes à vontade do próprio povo, considerada esta como a soberania popular.⁹⁰⁵ Sendo assim, as decisões são tomadas com base em pré-decisões⁹⁰⁶, como no caso das decisões judiciais que são feitas com base em escolhas constitucionais. Dessa forma, o procedimento como pré-compromisso, embora possa ser alterado, garante a existência de uma deliberação⁹⁰⁷ e que tal deliberação seja institucionalmente observada pelo sistema de IA judicial.

Nesse sentido, as instituições democráticas não devem ser estabelecidas como uma antiga forma de majoritarismo, mas sim com uma forma particular de decisão majoritária, que é baseada em princípios democráticos que observam a proteção às minorias⁹⁰⁸, pois, embora existam questões sobre as quais as pessoas discordam do ponto de vista razoável⁹⁰⁹, por outro lado se caracterizam questões de discordância que passam por um viés de discriminação.⁹¹⁰

Desse modo, a existência de divergências razoáveis é característica de todas as democracias modernas. Justamente em decorrência da própria característica do regime democrático, os juízes discordam tanto entre si quanto quaisquer outras

⁹⁰³ CONSELHO NACIONAL DE JUSTIÇA. Resolução nº 332, de 21 de agosto de 2020.

⁹⁰⁴ ELSTER, Jon. Ulysses and the sirens: A theory of imperfect rationality.

⁹⁰⁵ WALDRON, Jeremy. Op. cit., p. 275.

⁹⁰⁶ Ibid., p. 278.

⁹⁰⁷ Id.

⁹⁰⁸ Ibid., p. 279.

⁹⁰⁹ “As pessoas são razoáveis em um aspecto básico quando, entre iguais, por exemplo, estão dispostas a propor princípios e critérios como termos equitativos de cooperação e a submeter-se voluntariamente a eles, dada a garantia de que os outros farão o mesmo. Entendem que essas normas são razoáveis a todos e, por isso, as consideram justificáveis para todos, dispondo-se discutir os termos equitativos que outros propuseram. O razoável é um elemento da ideia de sociedade como um sistema de cooperação equitativa, e, que seus termos equitativos sejam razoáveis à aceitação de todos, faz parte da ideia de reciprocidade. Como já disse (I:3.2), a ideia de reciprocidade encontra-se entre a ideia de imparcialidade, que é altruísta (o bem geral constitui a motivação), e a ideia de benefício mútuo, compreendido como benefício geral com respeito à situação presente e futura, sendo as coisas como são” (RAWLS, John. **O Liberalismo Político**, p. 93).

⁹¹⁰ WALDRON, Jeremy. Op. cit., p. 279.

peçoas. Essa discordância, sobre a natureza da Constituição e sobre sua interpretação, é de boa-fé⁹¹¹, ainda que seja baseada em vieses equivocados.⁹¹² Os desacordos decorrem, inclusive, de imperfeição ou assimetria de informação⁹¹³, o que demonstra a necessidade de adequada delineação dos pré-compromissos que podem atuar como balizas interpretativas para a decisão judicial auxiliada por Inteligência Artificial.

A partir destes pressupostos, evidencia-se que um pré-compromisso, embora seja racional em sua gênese, passa a estimular decisões intuitivas pelas instituições democráticas, que podem aprimorar a agilidade decisória dentro das atividades estatais, especialmente no que toca às decisões judiciais nascidas do processo comunicacional cooperativo (duplo canal de correção) entre os sistemas de Inteligência Artificial e o julgador humano (racional e independente) na aplicação do Direito.

Isso se evidencia na medida em que há a superação do pré-compromisso individual (heurística individual) mediante o reconhecimento de que os pré-compromissos constitucionais se dão no plano coletivo, na dinâmica viva e entrópica da própria democracia, a qual deve se manifestar nos sistemas de Inteligência Artificial judicial. A parábola da biblioteca teológica utilizada por Waldron é bastante didática, na medida em que a sociedade pode decidir rever aspectos do pacto democrático com o passar dos anos, sem que isso autorize o incentivo a decisionismos através de sistemas de IA judicial.

3.4.2 Heurísticas e Pré-Compromissos: Heurísticas do Julgador *versus* Heurísticas do Direito

Do ponto de vista da psicologia, uma heurística pode ser definida como um atalho mental para uma decisão mais rápida⁹¹⁴, com grande nível de automatização e já reprodutora de um padrão pré-definido, ou mesmo como uma decisão que reverbere o padrão já conhecido pela mente humana. Daniel Kahneman diz que

⁹¹¹ Ibid., p. 280.

⁹¹² KAHNEMAN, Daniel. Op. cit., posição 570.

⁹¹³ DOWNS, Anthony. Teoría económica de la acción política en una democracia. In: RUBIO, Albert Batlle (coord.). **Diez textos básicos de Ciencia Política**. Madrid: Ariel, 2001, p. 100.

⁹¹⁴ MORAES, José Diniz de; TABAK, Benjamin Miranda. As Heurísticas e Vieses da Decisão Judicial: análise econômico-comportamental do direito. **Revista Direito GV**, São Paulo, v. 14, n. 2, p. 618-653, set. 2018. DOI: 10.1590/2317-6172201824.

heurística é um procedimento simples que auxilia na busca de respostas (ainda que provavelmente imperfeitas), bem como que tais respostas têm a característica de adequação para questões difíceis.⁹¹⁵ Veja-se que o estudo mais influente sobre o tema foi produzido em decorrência da parceria entre Daniel Kahneman e Amos Tversky, pois Kahneman vem desenvolvendo sua pesquisa neste ponto há algumas décadas.

Para facilitar a compreensão, de maneira mais apegada à didática do que a um rigorismo, Kahneman⁹¹⁶ divide as formas de pensar entre o Sistema 1 e o Sistema 2⁹¹⁷ como se fossem dois personagens. Nesse ponto, o presente trabalho passa a adotar a classificação didática das formas de pensamento, na maneira proposta por Kahneman, pois a divisão de funções entre os Sistemas 1 e 2 é bastante eficiente, eis que otimiza o desempenho e diminui o esforço cognitivo.⁹¹⁸

Assim, o intuitivo Sistema 1 tem mais influência e é o autor silencioso de várias das escolhas feitas e dos julgamentos realizados pelo ser humano⁹¹⁹, sendo que o Sistema 1 nunca é desligado.⁹²⁰ Observa-se que o Sistema 1 atua de maneira continua na aplicação de soluções intuitivas pré-programadas. Tal abordagem pode ser aplicada aos pré-compromissos⁹²¹, na medida em que esses atuam num nível de pré-compreensão interpretativa com potencial de vinculação das decisões judiciais pela via argumentativa-institucional.

Dessa forma, a função do Sistema 1 é manter e atualizar o modelo de mundo a ser observado pelo intérprete⁹²², razão pela qual deve passar de um modelo de mundo pessoal e puramente subjetivado para um modelo de mundo de maior objetivação (o que também é inerente aos sistemas de Inteligência Artificial judicial), já que nesse último deve existir uma abordagem jurídica pautada numa versão ontológica e valorativa pressuposta pelos pré-compromissos (constitucionais e legais) a serem observados na decisão judicial auxiliada por Inteligência Artificial.⁹²³

⁹¹⁵ KAHNEMAN, Daniel; TVERSKY, Amos. **Judgement Under Uncertainty: Heuristics and Biases**. Cambridge: Cambridge University Press, 1982, posição 232. E-book Kindle.

⁹¹⁶ KAHNEMAN, Daniel. Op. cit., posição 229.

⁹¹⁷ Ibid., posição 281.

⁹¹⁸ Ibid., posição 469.

⁹¹⁹ Ibid., posição 282.

⁹²⁰ Ibid., posição 473.

⁹²¹ WALDRON, Jeremy. Op. cit., p. 258-260.

⁹²² KAHNEMAN, Daniel. Op. cit., posição 1502.

⁹²³ WALDRON, Jeremy. Op. cit., p. 280.

Nesse sentido, inclusive em observância ao argumento contramajoritário⁹²⁴, as relações de causalidade jurídica indicadas pelos pré-compromissos constitucionais preparam o julgador para pensar por meio de padrões inseridos dentro de seu próprio Sistema 1, pois ele reproduz os referidos padrões.⁹²⁵ Dessa forma, tal causalidade intencional consiste na pressuposição de que algo irá acontecer a partir do que se conhece⁹²⁶, sendo papel dos pré-compromissos delimitar aquilo que “se conhece” juridicamente, a partir da interação com outros pontos de vista que são externos ao Direito (mas ínsitos à pessoa do julgador). Desse modo, tentar negar a influência da pré-compreensão externa ao Direito na composição da decisão judicial acabaria sendo uma tentativa de encobrir a realidade, pois a pré-compreensão é também delineada por questões de assimetria informacional e de escolhas pessoais⁹²⁷, seja nas decisões tomadas pelos seres humanos ou nas decisões auxiliadas por máquinas programadas pelos humanos.⁹²⁸

Logo, a causalidade intencional é uma consequência da heurística da disponibilidade utilizada nas decisões judiciais⁹²⁹, pois na “dúvida, o Sistema 1 aposta numa solução, e a aposta é orientada pela experiência”.⁹³⁰ Por outro prisma, a heurística da disponibilidade (*availability heuristic*)⁹³¹ é bastante criticada por Nassim Taleb⁹³², porque pode causar uma decisão baseada em viés de predição daquilo que vai ocorrer no futuro, embora o futuro seja imprevisível devido à característica entrópica da realidade informacional. Isso lembra que o Direito tem compromisso com a previsibilidade de sua aplicação nas decisões judiciais, seja a partir do direito fundamental à segurança jurídica (art. 5º, *caput*, da Constituição Federal, art. 30, *caput*, do Decreto-Lei 4.657/1942, art. 489, §1º, IV, art. 926 e art. 927 do Código de Processo Civil e art. 5º da Resolução 332/2020 do CNJ⁹³³) ou porque há uma expectativa razoável dos cidadãos de que haja um adequado grau de objetividade nos julgamentos (vide o Capítulo 1). Esse papel que, como foi comentado, é desenvolvido

⁹²⁴ GARGARELLA, Roberto. **La Justicia Frente al Gobierno**. Quito: Corte Constitucional para el Período de Transición, 2011, p. 122.

⁹²⁵ KAHNEMAN, Daniel. Op. cit., posição 1609.

⁹²⁶ Ibid., posição 1617.

⁹²⁷ DOWNS, Anthony. Op. cit., p. 106.

⁹²⁸ SUNSTEIN, Cass R. Algorithms, Correcting Biases.

⁹²⁹ Id.

⁹³⁰ KAHNEMAN, Daniel. Op. cit., posição 1675.

⁹³¹ Ibid., posição 145.

⁹³² TALEB, Nassim N. Op. cit., p. 192.

⁹³³ CONSELHO NACIONAL DE JUSTIÇA. Resolução nº 332, de 21 de agosto de 2020.

numa abordagem bastante ligada à dinâmica de construção e reconstrução dos consensos e de como os pré-compromissos atuam como um amortecedor jurídico na reconstrução consensual em sociedade.

Todavia, vê-se que, ao se admitir a heurística da disponibilidade no processo de formação das decisões judiciais, o enfoque não está apenas sobre os pré-conceitos decorrentes da historicidade da linguagem na formação do Direito, pois, para além disso, a abordagem se volta ao modo como a linguagem (e seus pré-conceitos intersubjetivos) encontra limites na própria forma de pensar, inerente a todo ser humano. Por isso, existe a necessidade de serem estabelecidas heurísticas jurídico-argumentativas como anteparo às heurísticas psicológicas do julgador, solução que surge a partir da abordagem de pré-compromissos democráticos. Dessa forma, os pré-compromissos possuem aspecto político-jurídico que reconhece a dimensão da linguagem na sua formação⁹³⁴, mas que não impedem o reconhecimento da irrevogável humanidade do julgador, motivo pelo qual a abordagem dos pré-compromissos como heurísticas de embasamento dos sistemas de IA judicial é uma proposta de controle racional dos decisionismos mais sincera para com a realidade, pois não buscam idealizar a racionalização humana como “a priori” superior, uma vez que reconhecem os limites desta razão, o que compartilha estrutura com a ideia de racionalidade limitada e com a função heurística dos argumentos jurídicos.⁹³⁵

Desse modo, os pré-compromissos constituem parte da arquitetura de escolhas⁹³⁶ político-jurídica que demonstra o poder do Sistema 1⁹³⁷ no estímulo das decisões a serem tomadas por julgadores. Isso ocorre porque o pensamento intuitivo nunca é desligado, quanto mais se nega a influência psicológica nos julgamentos, mais se mascara o fato de que na verdade o Sistema 1 influencia até mesmo as mais cuidadosas decisões⁹³⁸, haja vista que o *input* intuitivo nunca cessa. Em suma, a justificação argumentativa de uma decisão, por si só, não escapa das heurísticas que por ela são encapsuladas, o que ganha força com a possibilidade de que esse

⁹³⁴ WALDRON, Jeremy. Op. cit., p. 278.

⁹³⁵ “A última das funções positivas da dogmática é a *heurística*. A dogmática contém uma série de modelos de solução, distinções e pontos de vista que não apareceriam se se tivesse que começar sempre novamente. Este instrumental é, evidentemente, útil, ainda que a decisão não esteja determinada. Sugerem-se perguntas e respostas que de outra maneira seriam impossíveis ou inalcançáveis” (ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 236).

⁹³⁶ THALER, Richard H.; SUNSTEIN, Cass R. Op. cit., posição 1542-1543.

⁹³⁷ KAHNEMAN, Daniel. Op. cit., posição 1788.

⁹³⁸ Ibid., posição 1805.

estímulo argumentativo ocorra numa via comunicacional que flui da IA judicial para o julgador humano, e vice-versa, a fim de se estimular a pretensão de correção⁹³⁹ no processo decisório.

No sentido aqui tratado, do ponto de vista da teoria da argumentação, Robert Alexy reconhece a função de estabilização⁹⁴⁰ e a função heurística⁹⁴¹ no uso de enunciados dogmáticos (que possuem características de consensualidade). Desse modo, quanto à extensão da justificação se pode focar no uso justificativo (aplicação de enunciado dogmático justificando-o com outros argumentos) ou no uso não-justificativo, que ocorre quando os argumentos dogmáticos podem ser utilizados mesmo que eles próprios não sejam também fundamentados, quando o enunciado dogmático não é posto em dúvida.⁹⁴² Nessa linha, até certo ponto o uso não-justificativo (heurístico) é obrigatório, pois nenhum interprete “pode fundamentar todos os enunciados dogmáticos em que se apoia para a discussão do seu problema”⁹⁴³, uma vez que ninguém pode fundamentar tudo simultaneamente.⁹⁴⁴ Nesse sentido, percebe-se que a própria teoria da argumentação jurídica utiliza mecanismos heurísticos⁹⁴⁵, pois, ainda que se busque a saturação das possibilidades de argumentos a serem utilizados numa justificação⁹⁴⁶, quando se trata de se decidir com base em enunciados dogmáticos o uso não-justificativo (heurístico) se apresenta como barreira intuitiva⁹⁴⁷ à justificação desnecessária. Sendo assim, leva-se em consideração a possibilidade do uso não-justificativo de proposições normativas (ou enunciados dogmáticos) como elementos básicos de programação de Inteligência Artificial judicial pautada em representação do conhecimento⁹⁴⁸, já que nem todas as premissas de julgamento democrático devem sofrer modificação sem controle humano consciente no próprio código de programação (conforme o escopo pensando para cada sistema de IA em específico).

⁹³⁹ ALEXY, Robert. **Conceito e Validade do Direito**, p. 154-155.

⁹⁴⁰ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 232.

⁹⁴¹ *Ibid.*, p. 236.

⁹⁴² *Ibid.*, p. 228.

⁹⁴³ *Id.*

⁹⁴⁴ *Id.*

⁹⁴⁵ ATIENZA, Manuel. *Op. cit.*, p. 210.

⁹⁴⁶ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 216.

⁹⁴⁷ KAHNEMAN, Daniel. *Op. Cit.*, posição 826.

⁹⁴⁸ “An ontology that supports representing that kind of CSK would be a very complex affair. This is an example of the knowledge representation bottleneck that has impeded AI & Law research” (ASHLEY, Kevin D. **Artificial Intelligence and Legal Analytics**, p. 188).

Portanto, os pré-compromissos devem ser utilizados para que o pensamento intuitivo (automático) dos julgadores se adeque aos protótipos e ao conjunto de modelos padrão⁹⁴⁹ democraticamente estabelecidos para a argumentação jurídica auxiliada pela Inteligência Artificial.

Por outro lado, a abordagem de pré-compromissos como heurísticas a serem aplicadas pode vir a ser insuficiente e causadora de vieses de julgamento⁹⁵⁰, o que implica na necessária adoção de controles racionais do Sistema 2⁹⁵¹ de maneira a diminuir o risco de erros intuitivos.⁹⁵² Nesse sentido, o julgador e o sistema legal (desenhado para a condução à decisão correta) devem criar caminhos deliberativos para que no caso de dúvida ou incerteza seja sempre acionado o reflexivo Sistema 2⁹⁵³, na pessoa do julgador humano que tem o controle sobre o sistema de IA judicial. Assim, o Sistema 2 pode ser acionado na hipótese de ser detectado evento que transgrida o modelo de mundo jurídico mantido pelo Sistema 1⁹⁵⁴, bem como o Sistema 2 igualmente tem a tarefa de realizar o monitoramento perene de seu próprio comportamento.⁹⁵⁵

Contudo, ressalta-se que mesmo o Sistema 2, racional e mais lento⁹⁵⁶, pode incorrer em vieses que representam vícios decisórios.⁹⁵⁷ A partir dessa visão, vale destacar que o controle racional exercido pelo próprio julgador é sempre limitado pela forma de funcionamento do pensamento humano, pois o Sistema 1 pode vir a influenciar a compreensão do problema pelo Sistema 2.⁹⁵⁸ Sendo assim, os vieses efetivamente são uma preocupação, uma vez que estes, muitas vezes, não podem ser evitados, haja vista que existem momentos em que o Sistema 2 não oferece quaisquer indícios sobre o erro.⁹⁵⁹

Percebe-se que há uma aposta da teoria do direito na decisão racional típica do Sistema 2, lenta e analítica, contudo, tal pretensão protagonismo do Sistema 2⁹⁶⁰ é uma questão problemática, pois, em relação ao pensamento reflexivo (e aos seus

⁹⁴⁹ KAHNEMAN, Daniel. Op. Cit., posição 826.

⁹⁵⁰ THALER, Richard H.; SUNSTEIN, Cass R. Op. cit., posição 495-496.

⁹⁵¹ KAHNEMAN, Daniel. Op. cit., posição 536.

⁹⁵² Ibid., posição 1657.

⁹⁵³ Ibid., posição 1679.

⁹⁵⁴ Ibid., posição 461.

⁹⁵⁵ Ibid., posição 464.

⁹⁵⁶ Ibid., posição 1870.

⁹⁵⁷ Ibid., posição 1703.

⁹⁵⁸ Ibid., posição 1575.

⁹⁵⁹ Ibid., posição 570.

⁹⁶⁰ Ibid., posição 415.

efeitos na justificação de decisões judiciais), pode-se dizer que ele atua mediante operações trabalhosas, destacando-se que há uma resistência do Sistema 2 em colocar mais esforço do que seja estritamente necessário.⁹⁶¹ Dessa forma, quando se trata de decisões humanas, evidencia-se que há uma dominância do Sistema 1 (intuitivo) sobre o Sistema 2 (reflexivo), pois, muitas vezes, as decisões que o Sistema 2 imagina ter tomado são na verdade dirigidas pelo Sistema 1 na condição de silencioso protagonista.⁹⁶²

Nesse passo, evidencia-se a importância de conhecer o funcionamento das heurísticas envolvidas no processo de tomada de decisão humana, uma vez que as mesmas se refletem nas decisões judiciais.⁹⁶³

Em relação aos tipos de heurísticas envolvidas, primeiramente, verifica-se que a heurística da representatividade ocorre quando o Sistema 1 reconhece relações mais simples de inferência⁹⁶⁴ mediante observação de características aparentes e se destaca na integração de informações sobre algo para possibilitar sua decisão intuitiva. Contudo, ao tratar da heurística da representatividade, o Sistema 1 não realiza processamento múltiplo de tópicos de uma só vez e não é eficiente ao utilizar estatística pura como base informacional.⁹⁶⁵ São exemplos da heurística da representatividade “quando alguém diz ‘Ela vai ganhar a eleição; você pode perceber que é uma vencedora’ ou ‘Ele não irá longe como acadêmico; tem tatuagens demais’”.⁹⁶⁶

Já a heurística do afeto é uma substituição quando a resposta dada a uma questão fácil (um sentimento a respeito de algo) é utilizada como resposta para uma pergunta bem mais difícil (o que se pensa a respeito do tema)⁹⁶⁷, a qual acaba por simplificar a ação humana por meio de uma construção mental de cunho emocional que aparenta ser mais ordenada (racional) do que se verifica na realidade.⁹⁶⁸ Dessa maneira, a maior objetividade da heurística dos pré-compromissos⁹⁶⁹ indica um mecanismo de diminuição da incidência⁹⁷⁰ da heurística

⁹⁶¹ Ibid., posição 616.

⁹⁶² Ibid., posição 617.

⁹⁶³ THALER, Richard H.; SUNSTEIN, Cass R. Op. cit., posição 3742-3743.

⁹⁶⁴ KAHNEMAN, Daniel. Op. cit., posição 2037.

⁹⁶⁵ Ibid., posição 738.

⁹⁶⁶ Ibid., posição 3081.

⁹⁶⁷ Ibid., posição 2842.

⁹⁶⁸ Ibid., posição 2865.

⁹⁶⁹ WALDRON, Jeremy. Op. cit., p. 259.

⁹⁷⁰ KAHNEMAN, Daniel. Op. cit., posição 3480.

do afeto, na qual as decisões são alinhadas por sentimentos de gostar ou não gostar, havendo pouca análise deliberativa ou mesmo raciocínio.⁹⁷¹

Também há a heurística da disponibilidade, que pode ser definida “como o processo de julgar a frequência segundo a ‘facilidade com que as ocorrências vêm à mente’”.⁹⁷² Assim, os pré-compromissos funcionam como uma espécie de heurística da disponibilidade advinda legitimamente do próprio sistema jurídico⁹⁷³, os quais têm o potencial de sobrescrever (no todo ou em parte) as heurísticas de disponibilidade puramente psicológicas⁹⁷⁴ e em desacordo com a realidade jurídica.

Assim, Kahneman ressalta a inevitável confusão da própria democracia na medida em que as heurísticas de disponibilidade e do afeto⁹⁷⁵, que baseiam as crenças das pessoas, são tendenciosas de maneira natural, ainda que, quase sempre, apontem na direção correta. Em suma, a essência das heurísticas intuitivas se dá pelo fato de que “quando confrontados com uma questão difícil, muitas vezes respondemos a uma mais fácil em lugar dela, normalmente sem perceber a substituição”.⁹⁷⁶

Sendo assim, para diminuir a incidência de vieses na decisão, Kahneman destaca que é adequado eliminar a redundância das fontes de informação⁹⁷⁷, de maneira a se buscar um incremento informacional advindo da pluralidade de fontes. Entretanto, embora o Sistema 2 (reflexivo) tenha a oportunidade de rejeitar as respostas intuitivas ou modificá-las (incorporando novas informações), muitas vezes, ele funciona com o menor esforço e acaba aceitando uma resposta heurística sem maior exame sobre a adequação da mesma⁹⁷⁸, o que reforça a necessidade de construção de uma arquitetura constitucional de escolhas⁹⁷⁹ pautada nos pré-compromissos democráticos como as heurísticas a serem seguidas, podendo tais pré-compromissos condicionar intuitivamente as decisões judiciais para os casos em que o raciocínio lógico-analítico (Sistema 2) se apresente deficitário.

A título de exemplo, pelos tribunais inferiores e pelos juízes de primeiro grau, de maneira direta e intuitiva (sem grande necessidade de carga

⁹⁷¹ Ibid., posição 264.

⁹⁷² Ibid., posição 2652.

⁹⁷³ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 236.

⁹⁷⁴ KAHNEMAN, Daniel. Op. cit., posição 2056.

⁹⁷⁵ Ibid., posição 2977.

⁹⁷⁶ Ibid., posição 272.

⁹⁷⁷ Ibid., posição 1776.

⁹⁷⁸ Ibid., posição 2085.

⁹⁷⁹ THALER, Richard H.; SUNSTEIN, Cass R. Op. cit., posição 4108-4111.

reflexiva/argumentativa), pode ser aplicada a heurística inerente à forma dada ao direito de crítica, decorrente da liberdade de manifestação do pensamento e da liberdade de imprensa (art. 5º, IV e IX, e art. 220 da Constituição Federal), cujos limites foram devidamente delineados na *ratio decidendi* simbiótica da ADPF 130⁹⁸⁰ e da ADI 4451⁹⁸¹ julgadas pelo Supremo Tribunal Federal. Outro exemplo de heurística democrática a basear decisões judiciais é a adotada quando a fixação de índice de correção monetária não corresponde à realidade do fenômeno inflacionário, o que viola o direito à propriedade (art. 5º, XXII, e art. 170, II, da Constituição Federal),

⁹⁸⁰ “O pensamento crítico é parte integrante da informação plena e fidedigna. O possível conteúdo socialmente útil da obra compensa eventuais excessos de estilo e da própria verve do autor. O exercício concreto da liberdade de imprensa assegura ao jornalista o direito de expender críticas a qualquer pessoa, ainda que em tom áspero ou contundente, especialmente contra as autoridades e os agentes do Estado. A crítica jornalística, pela sua relação de inerência com o interesse público, não é aprioristicamente suscetível de censura, mesmo que legislativa ou judicialmente intentada. O próprio das atividades de imprensa é operar como formadora de opinião pública, espaço natural do pensamento crítico e ‘real alternativa à versão oficial dos fatos’” (BRASIL. Supremo Tribunal Federal (Plenário). **Ação Direta de Inconstitucionalidade 130**. Não recepção da Lei de Imprensa, Lei 5.250/1967, pela Constituição Federal de 1998, em razão da incompatibilidade com a liberdade de manifestação do pensamento, a liberdade de informação e comunicação e a liberdade de imprensa. Brasília, DF: Supremo Tribunal Federal, [2009]. Relator: Min. Carlos Ayres Britto, 30 de abril de 2009. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=AC&docID=605411>. Acesso em: 21 abr. 2021).

⁹⁸¹ “2. A livre discussão, a ampla participação política e o princípio democrático estão interligados com a liberdade de expressão, tendo por objeto não somente a proteção de pensamentos e ideias, mas também opiniões, crenças, realização de juízo de valor e críticas a agentes públicos, no sentido de garantir a real participação dos cidadãos na vida coletiva”. (...) “5. O direito fundamental à liberdade de expressão não se direciona somente a proteger as opiniões supostamente verdadeiras, admiráveis ou convencionais, mas também aquelas que são duvidosas, exageradas, condenáveis, satíricas, humorísticas, bem como as não compartilhadas pelas majorias. Ressalte-se que, mesmo as declarações errôneas, estão sob a guarda dessa garantia constitucional” (BRASIL. Supremo Tribunal Federal (Plenário). **Ação Direta de Inconstitucionalidade 4451**. Inconstitucionalidade do art. 45, II e III, e §4º e §5º, da Lei 9.504/1997, Lei de Eleições, em razão da ofensa à liberdade de manifestação do pensamento, à liberdade de comunicação e expressão artística e ao livre exercício profissional de humoristas. Brasília, DF: Supremo Tribunal Federal, [2018]. Relator: Min. Alexandre de Moraes, 21 de junho de 2018. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=749287337>. Acesso em: 21 abr. 2021).

conforme delimitado pelo Supremo Tribunal Federal (ADI 4357⁹⁸², ADI 4425⁹⁸³ e ADI 4425 QO⁹⁸⁴).

Para além dos exemplos citados, evidencia-se que nem todo precedente do Supremo Tribunal Federal⁹⁸⁵ pode ser classificado como alinhado à Constituição Federal, pois mesmo na corte constitucional podem existir decisões direcionadas por vieses de julgamento, desconsiderando-se as heurísticas normativo-democráticas delimitadas pela própria Constituição Federal. Um exemplo que pode ser apresentado é o julgamento de repercussão geral que resultou na total supressão do direito de greve dos servidores públicos da carreira da polícia civil (ARE 654432⁹⁸⁶), equívoco que levou à desconsideração do previsto no art. 37, VII, da Constituição Federal mediante a simples equiparação entre servidores civis e militares para fins de direito de greve, substituindo-se a escolha do constituinte prevista na regra excepcional do

⁹⁸² “O direito fundamental de propriedade (CF, art. 5º, XXII) resta violado nas hipóteses em que a atualização monetária dos débitos fazendários inscritos em precatórios perfaz-se segundo o índice oficial de remuneração da caderneta de poupança, na medida em que este referencial é manifestamente incapaz de preservar o valor real do crédito de que é titular o cidadão. É que a inflação, fenômeno tipicamente econômico-monetário, mostra-se insuscetível de captação apriorística (ex ante), de modo que o meio escolhido pelo legislador constituinte (remuneração da caderneta de poupança) é inidôneo a promover o fim a que se destina (traduzir a inflação do período)” (BRASIL. Supremo Tribunal Federal (Plenário). **Ação Direta de Inconstitucionalidade 4357**. Inconstitucionalidade parcial da Emenda Constitucional 62/2009, em razão da ofensa ao direito à propriedade. Brasília, DF: Supremo Tribunal Federal, [2013]. Relator: Min. Carlos Ayres Britto. Relator para acórdão: Min. Luiz Fux, 14 de março de 2013. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=6812428>. Acesso em: 21 abr. 2021).

⁹⁸³ BRASIL. Supremo Tribunal Federal (Plenário). **Ação Direta de Inconstitucionalidade 4425**. Inconstitucionalidade parcial da Emenda Constitucional 62/2009, em razão da ofensa ao direito à propriedade, à isonomia, à efetividade da jurisdição, à coisa julgada material, à separação dos poderes, fixação de critérios atualização monetária para dívidas públicas e de critérios para pagamento de precatórios. Brasília, DF: Supremo Tribunal Federal, [2013]. Relator: Min. Luiz Fux: 14 de março de 2013. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=5067184>. Acesso em: 21 abr. 2021.

⁹⁸⁴ BRASIL. Supremo Tribunal Federal (Plenário). **Questão de Ordem na Ação Direta de Inconstitucionalidade 4425**. Modulação dos efeitos da declaração de inconstitucionalidade parcial da Emenda Constitucional 62/2009. Relator: Min. Luiz Fux: 25 de março de 2015. Brasília, DF: Supremo Tribunal Federal, [2015]. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=9016259>. Acesso em: 21 abr. 2021.

⁹⁸⁵ Que aqui é utilizado apenas como um exemplo.

⁹⁸⁶ “A carreira policial é o braço armado do Estado, responsável pela garantia da segurança interna, ordem pública e paz social. E o Estado não faz greve. O Estado em greve é anárquico. A Constituição Federal não permite” (BRASIL. Supremo Tribunal Federal (Plenário). **Agravo em Recurso Extraordinário 654432 (Tema 541 de Repercussão Geral)**. Fixa o Tema 451 de repercussão geral, mediante a interpretação teleológica dos art. 9º, §1º, art. 37, VII, e art. 144, da Constituição Federal, a fim de estabelecer proibição absoluta ao exercício do direito de greve a todos servidores públicos integrantes das carreiras de segurança pública. Brasília, DF: Supremo Tribunal Federal, [2017]. Relator: Min. Edson Fachin. Relator para acórdão: Min. Alexandre de Moraes, 05 de abril de 2017. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=14980135>. Acesso em: 22 abr. 2021).

art. 142, §3º, IV, da Constituição Federal. O referido julgamento deixou de aplicar a heurística da disponibilidade em sua função democrática, inclusive pelo fato do Supremo Tribunal Federal fundamentar a proibição de greve aos servidores civis da segurança pública mediante a aplicação do art. 9º, §1º, da Constituição Federal, o qual prevê que a “lei definirá os serviços ou atividades essenciais” sujeitos a regime jurídico especial mais restritivo ao direito de greve.

No exemplo acima (ARE 654432), a suposta racionalidade (Sistema 2) encontra respaldo na institucionalização formal da decisão e na sua força como precedente vinculante que deve ser seguido. Tal característica de estabilização não se questiona nesta obra.⁹⁸⁷ Contudo, do ponto de vista material, a referida decisão seguiu heurísticas do afeto (foi utilizado o argumento prático de que o estado de greve seria “anárquico”) e de representatividade (“policial” não tem direito de greve, seja este um servidor público militar ou um civil), o que causou a deliberação enviesada. Ou seja, negar as heurísticas do sistema normativo, ainda que mediante racionalidade justificatória, é um viés interpretativo. Assim, os vieses decisórios se fazem presentes, mesmo quando existe uma fundamentação racional reflexiva (Sistema 2), o que pode levar a julgamentos equivocados.

Em resumo, na adoção das heurísticas corretas os pré-compromissos são parte do repositório de ações que o julgador pode adotar dentro da arquitetura constitucional de escolhas⁹⁸⁸, uma vez que a inteligência “não é apenas a capacidade de raciocinar; é também a capacidade de encontrar material relevante na memória e mobilizar a atenção quando necessária”.⁹⁸⁹ Por essa razão, percebe-se que a vigilância contínua do Sistema 2 nem sempre é um bem, além de ser impraticável⁹⁹⁰, motivo pelo qual se pode observar a importância de que o julgador esteja ambientado nas heurísticas imanentes aos pré-compromissos. Nesse caminho, os pré-compromissos são uma forma de se dar celeridade, previsibilidade e segurança aos julgamentos, mediante a possibilidade de análise intuitiva do julgador, previamente delimitada pelos acordos democráticos, o que ganha importância no caso de julgamento em que haja uma multiplicidade de resultados possíveis em decorrência da estrutura linguística das

⁹⁸⁷ A sociedade pode, por meios institucionais e democráticos, questionar as decisões do Supremo Tribunal Federal. Um exemplo disso é a resposta legislativa a determinadas decisões que são tomadas. O tema é extenso (e juridicamente instigante), mas foge aos limites desta obra.

⁹⁸⁸ THALER, Richard H.; SUNSTEIN, Cass R. Op. cit., posição 3929.

⁹⁸⁹ KAHNEMAN, Daniel. Op. cit., posição 950.

⁹⁹⁰ Ibid., posição 572.

normas aplicadas⁹⁹¹ ou devido à margem de decisão que o sistema atribui à figura do juiz.⁹⁹²

Nesse ponto, observa-se que a decisão judicial é uma batalha entre heurísticas psicológicas (do julgador) e normativas (do sistema jurídico), pois sempre haverá (em maior ou menor grau) as heurísticas de representatividade, do afeto e de disponibilidade com potencial de direcionar a decisão judicial para fora do modelo democrático. Fundamentação analítica (art. 93, IX, da CF e art. 489, §1º, do CPC) e justificação racional⁹⁹³ não são garantia de ausência de vieses decisórios intuitivos do julgador e são, até mesmo, imperceptíveis ao mesmo. Dessa forma, a batalha na mente do julgador se dará entre a visão de mundo por ele mesmo imaginada (e enviesada em suas próprias heurísticas) e a heurística da disponibilidade inerente aos pré-compromissos democráticos.

Nesse caminho, consistente em relação às heurísticas democráticas, as regras normativas⁹⁹⁴ são o primeiro nível heurístico voltado à “promoção de previsibilidade, confiabilidade e certeza”⁹⁹⁵, sendo que, em níveis⁹⁹⁶ subsequentes⁹⁹⁷, as justificações de primeira camada (fim imediato que a regra promove) e de segunda camada (fundamentos inerentes à espécie de justiça que a regra promove) podem ser

⁹⁹¹ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 199-200.

⁹⁹² HART, Herbert L. A. Op. cit., p. 165.

⁹⁹³ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 187.

⁹⁹⁴ *“The excursus into and evaluation of rule-sensitive particularism provides greater confidence in both the conceptual possibility and the normative plausibility of a decisional mode in which decision-makers treat the result indicated by a rule as itself a reason for action, even when that result diverges from the result that would have been indicated by direct application of the justifications lying behind the rule, and even when that result diverges from the result that would have been reached, within a given substantive theory of decision, by direct application of all relevant considerations recognized within that substantive theory of decision. This strong sense of rule-based decision-making, therefore, requires not only that rules be taken at least partially to impede recourse to the justifications behind the rule, but also that evaluating the strength of the rule-generating justifications as they apply to the instant case not be left solely to the decision-maker”*. (SCHAUER, Frederick. Op. cit., p. 100).

⁹⁹⁵ *“If the virtues of rules are seen to reside largely in the fostering of predictability, reliability, and certainty, then rule-sensitive particularism appears to take account of these virtues without committing the sin of rule-worship”* (Ibid., p. 98).

⁹⁹⁶ *“Once we see that justifications exist on different levels, we can appreciate that what is a justification with respect to its instantiation is usually itself also the instantiation of a deeper justification”* (Ibid., p. 74).

⁹⁹⁷ *“Divorcing the idea of a rule from its most specific and canonized formulation makes it possible to appreciate the intriguing way in which there may be stacked layers of rules, and the way in which the concept of a rule may be seen to be more of a relationship than an isolated entity.*

Until now I have been operating with a model of decision-making involving only formulated or formulatable rules and the justifications that lie behind them. Those background justifications themselves, however, do not appear from thin air, nor can the kinds of justifications I have been considering be thought to be in any sense foundational or ultimate. Rather, these comparatively specific justifications (peace and quiet in the park; no dangerous driving; no annoyance of patrons in the restaurant) are themselves instantiations of deeper and yet more abstract justifications” (Ibid. p. 73).

acionadas para tratamento de exceções ou hipóteses de derrotabilidade⁹⁹⁸, através de proposições normativas programadas ou sugeridas por Inteligência Artificial judicial (que tem sua função heurística do processo de julgamento) ou, mediante pensamento reflexivo e justificado, pelo julgador humano. Essa dinâmica comunicacional cooperativa é inerente ao duplo canal de correção conceituado nesta obra (Capítulo 1), pois parte de propostas de decisões menos refletidas auxiliadas pela IA judicial (pautada em racionalidade limitada) para alcançar decisões humanas com maior nível de racionalidade que superem as primeiras, as quais também podem ser testadas pela IA (desde que o controle permaneça humano).

Portanto, existem duas funções heurísticas, típicas dos pré-compromissos, que consistem: a) no condicionamento do pensamento do julgador segundo o modelo da democracia constitucional; e b) na maior previsibilidade da aplicação das normas jurídicas nas decisões judiciais, diminuindo os sobressaltos interpretativos.

Assim, por ser baseado na interação entre pensamento intuitivo (com suas heurísticas e vieses) e pensamento reflexivo, o processo decisório deve ser submetido aos controles heurístico-normativos da democracia (pré-compromissos), o que, se bem planejado e adequadamente executado, pode ser aprimorado pelos sistemas de IA judicial. Desse modo, podem ser definidos controles democráticos necessários e mais conscientes do processo de formação da decisão judicial.

Torna-se necessário o reconhecimento das funções das heurísticas dos pré-compromissos (e dos consensos democraticamente tensionados a eles antecedentes), seja como limites à justificação do pensamento do julgador ou como bases de disponibilidade de fundamentação de novas decisões. Por isso, a heurística tem efeitos a ela inerentes, não se apegando aos limites cognitivos que supostamente teriam os juízes.

Logo, saber acerca do funcionamento das heurísticas é fundamental para que se possa reconhecer uma decisão judicial equivocada dos pontos de vista racional e democrático, pois, mesmo uma decisão intuitiva pode receber um invólucro de justificação racional no momento de sua elaboração. Um exemplo disso é o *efeito halo*⁹⁹⁹, no qual, a partir de uma informação real (justificada em razão prática ou normativamente, no contexto da presente obra), o julgador da situação pressupõe

⁹⁹⁸ Ibid., p. 76.

⁹⁹⁹ KAHNEMAN, Daniel. Op. cit., posição 1705-1779.

intuitivamente diversas outras características que são comumente correlacionadas com a primeira, embora não haja comprovação destas últimas. Em casos em que o comportamento de partes é analisado (ações indenizatórias, litigância de má-fé e infrações de trânsito, por exemplo¹⁰⁰⁰), o *efeito halo* pode enviesar o julgamento.

Tais considerações indicam que o modelo de interação entre justificação racional das decisões e teoria dos precedentes, ainda que tenha realizado reconhecidas contribuições para o controle democrático das decisões judiciais, por si só não explica todos os aspectos do fenômeno decisório. Desse modo, o vácuo deixado entre a teoria da justificação racional e a teoria dos precedentes pode ser preenchido pelo duplo canal de correção (sistemas de IA judicial e julgador humano), que se utiliza da funcionalidade teórica dos pré-compromissos democráticos e das heurísticas de disponibilidade a eles inerentes.

Assim, as funções heurísticas dos pré-compromissos têm o potencial de aprimorar a vinculação do pensamento do julgador ao modelo da democracia constitucional e de estimular uma maior previsibilidade das decisões judiciais auxiliadas por Inteligência Artificial.

¹⁰⁰⁰ Obviamente, os exemplos de réus em processos penais são ainda mais problemáticos, mas frente às limitações de análise probatória penal impostas pela Res. 332/2020 do CNJ, demonstra-se adequado expor aqui outros tipos de casos.

4 JUSTIFICAÇÃO DA DECISÃO JUDICIAL E INTELIGÊNCIA ARTIFICIAL

4.1 O Juiz Ciborgue e a Pretensão de Correção: Facetas da Interação entre Suporte Fático, Fatores (“*Factors*”) e Conceitos Jurídicos Intermediados

4.1.1 O Suporte Fático Percebido pela Inteligência Artificial

O senso comum jurídico¹⁰⁰¹, via de regra, endereça o debate sobre a aplicação de Inteligência Artificial no Direito para o aspecto normativo. Contudo, muitas vezes, o debate sobre a lógica normativa, o alcance de regras e princípios e a força de representação da realidade atrelada a cada espécie de norma se esquece que há um universo antecedente, que é o mundo fático, ou o mundo da vida, o tecido onde os atos e fatos se desenrolam no espaço-tempo. Dessa forma, para que seja possível o duplo canal de correção e a explicabilidade a ele inerente, deve existir consistência normativa na aplicação do direito com o auxílio de sistemas de Inteligência Artificial, caminho por onde transita o presente estudo.

Na busca de explicar a relação entre o mundo da vida e o mundo normativo, Alchourrón e Bulygin fazem uma integração entre os casos e os seus elementos fáticos e as normas e suas proposições normativas.¹⁰⁰² Entre as normas jurídicas, as proposições normativas e o Universo de Casos há uma correlação advinda da manifestação de determinados fatos, atos e estados que correspondem a um caso, que, por sua vez, será relacionada às normas e levará a uma proposição normativa. A abordagem de Alchourrón e Bulygin tem sua relevância ao explicar a composição do Universo dos Casos (UC) através dos fatos, o que tem especial valia para explicitar o funcionamento de sistemas de IA judicial pautados no modelo de raciocínio baseado em casos (*case based reasoning*), por exemplo.

¹⁰⁰¹ WARAT, Luís Alberto. Saber Crítico e Senso Comum Teórico dos Juristas.

¹⁰⁰² BULYGIN, Eugenio. **Lógica deóntica, normas y proposiciones normativas**, p. 74

Desse modo, o Universo de Discursos (UD)¹⁰⁰³ se manifesta através de propriedades¹⁰⁰⁴, que formam o Universo de Propriedades (UP)¹⁰⁰⁵, sendo que cada propriedade representa certas circunstâncias relevantes¹⁰⁰⁶ através de atos, fatos ou estados. Essas propriedades¹⁰⁰⁷, que, por sua vez, representam um conjunto finito¹⁰⁰⁸, influenciam na construção do Universo de Casos (UC).¹⁰⁰⁹ A partir dos discursos, das propriedades e dos casos, no pensamento de Bulygin e Alchourrón, a passagem para o mundo normativo ocorre através do Universo de Ações (UA)¹⁰¹⁰, as quais podem ser descritas através de modais deônticos de obrigação, proibição, permissão ou facultatividade de determinada ação.¹⁰¹¹ As ações representadas em relação a

¹⁰⁰³ *"In order to answer this question we must know whether a certain (restitution of real state) is obligatory or not. So we are concerned with what we may conveniently call deontic status of action. There is a certain set of situations or states of affairs in which this action may take place and this we shall call a Universe of Discourse (UD)"* (ALCHOURRÓN, Carlos E.; BULYGIN, Eugenio. Op. cit., p. 10).

¹⁰⁰⁴ *"The states of affairs which belong to a UD are the elements of UD. All elements of UD share certain property which is the defining property of the UD. The Universe of Discourse may be described therefore as the set of all elements (states of affairs) identified by a certain property"* (Id.).

¹⁰⁰⁵ *"Any set of properties which may be present or absent in the elements of a UD we shall call a Universe of Properties (UP)"* (Ibid., p. 12).

¹⁰⁰⁶ *"The selection of the relevant circumstances or properties is therefore a value problem. If certain properties seem to be quite irrelevant for our problem (such as, e. g., the colour of the owner or the size of his nose) this is so only because there is a more or less unanimous consensus (in a given social group) concerning certain values. But this is, merely, a contingent fact; there is nothing necessary about it"* (Ibid., p. 11).

¹⁰⁰⁷ *"El ámbito fáctico del problema está constituido por las propiedades de UP y por los compuestos veritativo-funcionales de tales propiedades (siempre que no sean tautológicos o contradictorios); ellas determinan en un caso posible. Si la propiedad definitoria de un caso es una conjunción que contiene cada propiedad de UP o su negación, pero no ambas, entonces el caso es elemental y el conjunto de todos los casos elementales correspondientes a un UP se llama Universo de Casos (UC). El número de todos los casos elementales posibles correspondientes a un UP puede ser calculado fácilmente: es 2^n , donde 'n' es el número de propiedades de UP. Así, UC es el conjunto de todas las posibles situaciones fácticas"* (BULYGIN, Eugenio. **Lógica deôntica, normas y proposiciones normativas**, p. 51).

¹⁰⁰⁸ *"Sin embargo, no todo Universo de Casos tiene su origen en un conjunto finito de propiedades seleccionadas como relevantes. Es relativamente frecuente en la práctica jurídica caracterizar a los casos mediante valores numéricos, por ejemplo, en materia impositiva. En tales situaciones un UC puede contener un infinito, o al menos ilimitados, número de casos. Este hecho hace conveniente dar una definición general de UC: cualquier conjunto de propiedades que forme una división, i. e. una partición exhaustiva de UD, mutuamente excluyente e que no sea lógicamente vacía, es un UC"* (Id.).

¹⁰⁰⁹ *"The set of all elementary cases (corresponding to a UP) will be called a Universe of Cases (UC). The number of all possible elementary cases may be easily determined. Let n be the number of the properties of the UP; then 2^n is the number of the elementary cases, i. e. of the elements of the corresponding UC. The notion of a UC is, together with the number of the elementary cases, relative to a UP"* (ALCHOURRÓN, Carlos E.; BULYGIN, Eugenio. Op. cit., p. 12).

¹⁰¹⁰ *"A normative problem may be regarded as a question concerning the deontic status of certain actions, i. e. whether these actions are permitted or prohibited or obligatory, etc. Some of these actions are basic in the sense the other actions are truth-functional compounds of them. Any finite set of such basic actions will be called a Universe of Actions (UA)"* (Ibid., p. 10).

¹⁰¹¹ *"Solutions have been described in ch. I as deontic modalizations of the elements of the UA and their truth-functional compounds. The elements of the UA are generic actions. The distinction between generic and individual actions is analogous to the distinction between generic and individual cases (ch. II, sect. 3). The solutions that concern us in this context are therefore generic, not individual solutions."*

determinado microssistema têm independência lógica¹⁰¹² quanto às propriedades que levam à inferência. Nesse caminho, as possíveis resoluções¹⁰¹³ mediadas pelo sistema normativo pertencem ao Universo de Soluções (US).¹⁰¹⁴ O modelo de Bulygin tem abordagem que reconhece os potenciais e os limites da lógica de proposições normativas em relação às lacunas valorativas e às lacunas normativas¹⁰¹⁵ de cada

As symbols for the elements of the UA we shall use the letters p, q, r, etc. These letters may be interpreted as standing for propositions which describe generic actions, or better; states of affairs which are the result of an action" (Ibid., p. 35).

¹⁰¹² *"It will be assumed that the elements of the UA fulfil the following conditions:*

a) They are locally independent. This implies the acceptance of the hypothesis of logical atomism (cf. ch. II, sect. 1) regarding the elements of the UA. Were we to abandon the requirement of logical independence, it would be necessary to introduce meaning postulates, but by adopting the hypothesis of logical atomism we avoid this complication.

b) The elements of the UA are logically independent of the properties of the UP (cf. ch. II, sect. 1).

To assume a) and b) is equivalent to assuming the logical independence of the set formed by the properties of the UP and the elements of the UA" (Ibid., p. 35-36).

¹⁰¹³ *"De igual manera, podemos calcular todas las posibles soluciones o respuestas posibles al problema. Como nuestro problema es normativo, se refiere al status normativo de ciertas acciones que pueden ser deónticamente modalizadas como obligatorias, prohibidas o permitidas. Algunas de esas acciones son básicas en el sentido de que otras acciones son compuestos veritativo-funcionales de ellas. Un conjunto de tales acciones (siempre que no sean tautológicos o contradictorios) es llamado contenido deóntico. Por soluciones se entiende a cualquier contenido deóntico precedido por un operador deóntico (obligatorio, prohibido, permitido o facultativo) y sus compuestos veritativo-funcionales (siempre que no sean tautológicos ni contradictorios). Si la solución determina todos los contenidos deónticos de UA, se llama solución maximal. El conjunto de todas las soluciones maximales relativas a UA es el Universo de Soluciones Maximales (USmax)" (BULYGIN, Eugenio. **Lógica deóntica, normas y proposiciones normativas**, p. 52).*

¹⁰¹⁴ *"Un problema normativo si consiste en esclarecer si un cierto conjunto de acciones o una acción son obligatorias, prohibidas o permitidas. Cinco universos son usados en este análisis: Universo del Discurso (UD), Universo de Propiedades (UP), Universo de Casos (UC), Universo de Acciones (UA) y Universo de Soluciones (US).*

El Universo del Discurso es el conjunto de todas las situaciones particulares en las que la acción (o acciones) en cuestión puede(n) ser ejecutada(s). El Universo de Propiedades es el conjunto de todas las propiedades que tienen que ser tenidas en cuenta para la solución del problema; esas propiedades pueden estar presentes o ausentes en los elementos o situaciones de UD. Qué propiedades son consideradas relevantes en un problema valorativo, pero normalmente ese problema es (al menos parcialmente) resuelto por las normas formuladas por el legislador, o por precedentes (de decisiones judiciales previas). Así, tenemos que encontrar aquello que las normas jurídicas (leyes, precedentes, o normas consuetudinarias) prescriben para este problema" (Ibid., p. 50-51).

¹⁰¹⁵ *"Dado que las propiedades de UP usualmente no son numerosas (rara vez son más de cuatro o cinco) el número de casos elementales será, en la práctica, más bien limitado: muy raramente más de dieciséis o treinta y dos. Esto muestra que es relativamente fácil determinar cuáles son todas las posibles circunstancias fácticas (casos). Por ello, las preocupaciones de los juristas acerca de que el legislador no puede prever todas las posibles circunstancias futuras es absolutamente injustificada: no hay necesidad de que el legislador prevea nada que su tarea no es prever, sino decidir. Él decide (sobre la base de valoraciones) qué propiedades son relevantes para el problema y así clasifica a todos los elementos de UD. Todas las otras propiedades son simplemente irrelevantes conforme a su decisión. Por supuesto, puede ocurrir que las valoraciones de los jueces no coincidan con las del legislador: una propiedad que de acuerdo con las normas es irrelevante puede ser considerada como relevante por el juez y viceversa: el juez puede considerar que una propiedad que es relevante para el legislador no debería ser así considerada. Esas clases de problemas dan lugar a lo que se denominan 'lagunas axiológicas', que son muy diferentes a las lagunas normativas" (Ibid., p. 51-52).*

microssistema¹⁰¹⁶ jurídico¹⁰¹⁷, o que admite a entropia jurídica inerente e contingente ao duplo canal de correção, que é necessário no modelo de juiz ciborgue.

Além disso, o modelo de pensamento desenvolvido por Bulygin e Alchourrón é compatível com a Programação Orientada a Objetos (POO)¹⁰¹⁸, principalmente quando consideradas as diferentes classes de objetos e a “herança”¹⁰¹⁹ de atributos e de métodos¹⁰²⁰/funções¹⁰²¹ de uma classe em relação a outra classe num programa. De certa forma, falar em Universo de Discursos (UD), Universo de Propriedades (UP), Universo de Casos (UC), Universo de Ações (UA) e Universo de Soluções (US), em um microssistema finito e controlado, é pensar um modelo de raciocínio jurídico-computacional que leve em conta a herança de atributos e funções entre diferentes classes.

¹⁰¹⁶ “Es importante remarcar que: primero, un sistema normativo puede contener no solamente normas sino también otras clases de enunciados, e. g. definiciones y reglas conceptuales en general, al segundo, que los enunciados que constituyen la base del sistema pueden ser de diferentes orígenes, tipos y número. Ello convierte a la noción de sistema normativo en algo mucho más flexible, ya que no está ligada a una teoría jurídica de ideología política” (Ibid., p. 53).

¹⁰¹⁷ “Por un sistema deductivo se entiende a cualquier conjunto de enunciados que incluye todas sus consecuencias lógicas y, si entre sus consecuencias hay normas, entonces es un sistema normativo” (Ibid., p. 52-53).

¹⁰¹⁸ “A programação orientada a objetos pegou as melhores ideias da programação estruturada e combinou-as com vários conceitos novos. O resultado foi uma maneira diferente de organizar um programa. De um modo mais geral, um programa pode ser organizado de uma entre duas maneiras: a partir de seu código (o que está ocorrendo) ou a partir de seus dados (o que está sendo afetado). Com o uso somente da programação estruturada, normalmente os programas são organizados a partir do código. Essa abordagem pode ser considerada ‘o código atuando sobre os dados’.

Os programas orientados a objetos funcionam ao contrário. São organizados a partir dos dados, com o seguinte princípio-chave: ‘dados controlando o acesso ao código’. Em uma linguagem orientada a objetos, você define os dados e as rotinas que podem atuar sobre eles. Logo, um tipo de dado define precisamente que tipo de operações podem ser aplicadas a esses dados” (SCHILDT, Herbert. **Java para iniciantes**: crie, compile e execute programas Java rapidamente. Tradução: Aldir José Coelho Corrêa da Silva. 6ª ed. Porto Alegre: Bookman, 2015, p. 9).

¹⁰¹⁹ “Herança é o processo pelo qual um objeto pode adquirir as propriedades de outro objeto. Isso é importante porque dá suporte ao conceito de classificação hierárquica. Se você pensar bem, grande parte do conhecimento pode ser generalizada em classificações hierárquicas (isto é, top-down). Por exemplo, uma maçã Red Delicious faz parte da classificação *maçã*, que por sua vez faz parte da classe *fruta*, que fica sob a classe maior *alimento*. Isto é, a classe *alimento* possui certas qualidades (comestível, nutritivo, etc.) que, logicamente, também se aplicam à sua subclasse, *fruta*. Além dessas qualidades, a classe *fruta* tem características específicas (suculenta, doce, etc.) que a distingue de outros alimentos. A classe *maçã* define as qualidades específicas de uma maçã (cresce em árvores, não é tropical, etc.). Por sua vez, uma maçã Red Delicious herdaria as qualidades de todas as classes precedentes e só definiria qualidades que a tornam única.

Sem o uso de hierarquias, cada objeto teria que definir explicitamente todas as suas características. Com o uso da herança, um objeto só tem que definir as qualidades que o tornam único dentro de sua classe. Ele pode herdar seus atributos gerais de seu pai. Logo, é o mecanismo de herança que possibilita um objeto ser uma instância específica de um caso mais geral” (Ibid., p. 11).

¹⁰²⁰ Em Java, por exemplo, os métodos são as rotinas que podem ser demandadas de um programa para a sua execução.

¹⁰²¹ Em determinadas linguagens de programação, como em Python, por exemplo, as ações que podem ser chamadas à execução num programa são denominadas como funções.

A partir dessa compreensão acerca da necessária relação entre o mundo da vida e o mundo normativo, quando se pensa na teoria da norma aplicada à IA, o debate pode ser compreendido através do estudo do **suporte fático** atrelado ao aspecto normativo decorrente do **bem protegido** e da **intervenção**¹⁰²² em determinado direito¹⁰²³, pois o suporte fático é necessário para a compreensão do problema inerente à representação (ainda que parcial) do Direito em sistemas de IA judicial.

A interação entre mundo da vida (ontológico) e mundo normativo se dá através do suporte fático, sendo que este pode ser classificado em suporte fático abstrato e suporte fático concreto. Desse modo, o **suporte fático abstrato** se verifica nos atos, fatos ou estados previstos na norma jurídica e que, caso sejam realizados, levam à consequência jurídica¹⁰²⁴. Já o **suporte fático concreto** é a ocorrência, em si, dos atos, fatos ou estados previstos na norma jurídica.¹⁰²⁵

Nesse sentido, “a consequência jurídica de uma norma é produzida quando todas as suas condições são satisfeitas”¹⁰²⁶ (o que se verifica através da ocorrência, ou não, das propriedades que representam o caso), e somente pode ser delimitada

¹⁰²² “Embora normalmente sem referência à expressão ‘suporte fático’ ou a alguma teoria sobre ele, é comum que se pergunte se esse ou aquele ato, fato ou estado é protegido por essa ou aquela norma que garante um direito fundamental, ou se essa ou aquela ação estatal configura, ou não, uma intervenção nesse âmbito de proteção” (AFONSO DA SILVA, Virgílio. **Direitos Fundamentais**: conteúdo essencial, restrições e eficácia. São Paulo: Malheiros, 2009, p. 68).

¹⁰²³ “A construção do suporte fático composto pelo bem protegido e pela intervenção baseia-se nas relações entre os conceitos de intervenção, de bem protegido, de restrição e de proibição definitiva e *prima facie* (direito definitivo e *prima facie*), as quais podem ser expressas por meio de duas leis se seguem – denominadas *leis de intervenção* –, bem como por meio de suas contraposições:

- (I) Todas as medidas que sejam intervenções em um bem protegido por um direito fundamental são *prima facie* proibidas pelo direito fundamental.
- (II) Todas as medidas que sejam intervenções em um bem protegido por um direito fundamental, e que não sejam justificadas por uma restrição, são definitivamente proibidas pelo direito fundamental.
- (I') Todas as medidas que não sejam *prima facie* proibidas pelos direitos fundamentais não constituem intervenções em um bem protegido por um direito fundamental.
- (II') Todas as medidas que não sejam definitivamente proibidas pelos direitos fundamentais ou não constituem intervenções em um bem protegido por um direito fundamental ou estão justificadas por meio de uma restrição.

Essas leis fazem com que fique claro que os conceitos de bem protegido e de intervenção não têm necessariamente que determinar os primeiros passos sistemáticos em um controle de constitucionalidade no âmbito dos direitos fundamentais” (ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 306-307).

¹⁰²⁴ “Suporte fático *abstrato* é o formado, em linhas ainda gerais, por aqueles fatos ou atos do mundo que são descritos por determinada norma e para cuja realização ou ocorrência se prevê determinada consequência jurídica: preenchido o suporte fático, ativa-se a consequência jurídica” (AFONSO DA SILVA, Virgílio. **Direitos Fundamentais**, p. 67-68).

¹⁰²⁵ “*Suporte fático concreto*, intimamente ligado ao abstrato, é a ocorrência concreta, no mundo da vida, dos fatos ou atos que a norma jurídica, em abstrato juridicizou” (Ibid., p. 68).

¹⁰²⁶ ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 308.

através do suporte fático. Com base nisso, Alexy faz a distinção entre suporte fático em sentido amplo e suporte fático em sentido estrito.

O **suporte fático em sentido amplo** leva à estrutura triádica de bem protegido, intervenção e restrição no direito fundamental, de maneira que o conceito de suporte fático se coloca contraposto ao conceito de restrição, pois esta dinâmica é inerente ao caráter *prima facie* das normas jurídicas¹⁰²⁷, especialmente das que possuem estrutura de princípio.

Essa construção triádica do suporte fático em sentido amplo supera, em capacidade descritiva e em plasticidade normativa, a simples contraposição dual¹⁰²⁸ entre bem protegido e proibição de qualquer intervenção, ainda que essa última seja uma lógica passível de ser adotada em sistemas de IA judicial mais rudimentares.¹⁰²⁹ Entretanto, a construção dual da relação entre intervenção e suporte fático não é recomendável, porque, nessa hipótese, o suporte fático deixa de ser o contraponto ao conceito de restrição de Direito, de maneira que “com a satisfação do suporte fático, e caso não ocorra nenhuma restrição, o direito *prima facie* ou a proibição *prima facie* tornam-se direito definitivo ou proibição definitiva”.¹⁰³⁰

Em suma, numa leitura do pensamento de Alexy¹⁰³¹, a construção triádica entre suporte fático, intervenção e restrição leva à configuração de uma construção argumentativa que pode ser assim expressa:

¹⁰²⁷ “Ora, nada há que milite contrariamente à estrutura triádica (bem protegido/intervenção/restrrição); pelo contrário, ela é até mesmo sugerida em (5). Mas isso ainda não é uma razão suficiente a favor de uma construção do suporte fático somente com base no interesse protegido. **A ideia do conceito de suporte fático como um conceito contraposto ao conceito de restrição tem como tarefa abranger todos os pressupostos materiais para a ocorrência da consequência jurídica *prima facie***” (Ibid., p. 305).

¹⁰²⁸ “É possível dar às normas que garantem direitos a ações negativas uma formulação na qual aparece na cláusula condicional apenas o bem protegido. Um exemplo seria:

(2) Se uma ação constitui uma escolha profissional, então, há um direito *prima facie* a que nela não se intervenha.

A formulação equivalente para o dever correspondente seria:

(3) Se uma ação constitui uma escolha profissional, então, intervenções nessa ação são *prima facie* proibidas.

A vantagem dessa construção é que a partir dela se obtém um suporte fático bastante simples. O suporte fático simplesmente descreve o bem protegido”. (...) “O suporte fático dessa construção pode ser designado como ‘suporte fático composto pelo bem protegido’” (Ibid., p. 304).

¹⁰²⁹ Mais uma vez cabe ressaltar que o intérprete humano é o responsável pela interpretação, sendo que o pensamento humano será mais demandado caso se opte pela adoção de uma estrutura lógica mais simples (dual) de relação entre suporte fático e intervenção em um Direito.

¹⁰³⁰ ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 304.

¹⁰³¹ “Se se escolhe o suporte fático do bem protegido, o dualismo entre suporte fático e restrição é substituído por um trialismo entre suporte fático, intervenção e restrição. Uma análise mais detida demonstra, no entanto, que isso nada mais significa que a divisão do suporte fático em dois elementos. Assim, (3) pode ser reformulado como:

- (1) Se uma conduta integra o suporte fático, então, ela é *prima facie* permitida.
- (2) Se uma ação é protegida pelo bem jurídico, então, há um direito *prima facie* a que nela não se intervenha.
- (3) Se uma ação constitui um fato, ato ou estado protegido pelo bem jurídico, então, intervenções nessa ação são *prima facie* proibidas.
- (4) Se uma medida é uma intervenção no bem protegido, então, essa medida é *prima facie* proibida.
- (5) Se uma ação *h*, realizada por um titular de um direito é um ato ou fato que decorre do bem protegido, e se a medida *m* intervém na ação *h*, então *m* é *prima facie* proibida.¹⁰³²

Dessa maneira, verifica-se um “suporte fático composto pelo bem protegido e pela intervenção”¹⁰³³, o que pode ser aproveitado para a construção de sistemas de IA judicial voltados ao auxílio ao julgamento.

Por outro lado, vê-se que a via linguística de interlocução entre o mundo da vida e a norma jurídica é, muitas vezes, tratada com bastante simplicidade pelos profissionais do Direito, pois esses acionam suas ontologias e heurísticas mentais e, com relativa facilidade, conseguem propor aplicação para as normas jurídicas. As dificuldades de aplicação variam, aos intérpretes humanos, conforme a multiplicidade de significados e de proposições se avoluma (desde regras de direito sancionador com interpretação estrita até direitos fundamentais com estrutura de princípio).

Nesse sentido, se, por um lado, as normas representam um desafio em termos de lógica de programação jurídica, por outro lado, o suporte fático indica o tamanho monumental que é o desafio de fazer o sistema judicial inteligente perceber os fatos.

Quando se pensa no suporte fático em sentido amplo, o qual leva em conta os conceitos de suporte fático abstrato e concreto, o âmbito de justificação externa assume o papel de “fundamentação das premissas utilizadas na argumentação

(4) Se uma medida é uma intervenção na escolha profissional, então, essa medida é *prima facie* proibida.

E a (4) se pode dar a seguinte forma:

(5) Se uma ação *h*, realizada por um titular de um direito fundamental é uma escolha profissional, e se a medida *m* intervém em *h*, então *m* é *prima facie* proibida”

(Id.).

¹⁰³² Esta obra usará os símbolos utilizados por Alexy, embora as letras remetam a palavras da língua alemã e possam, a gosto do leitor, ser substituídas por letras que representem palavras na língua portuguesa, sem que isso leve à perda da compreensão.

¹⁰³³ ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 305.

interna”¹⁰³⁴, pois atua através de premissas consistentes em: “(1) regras de direito positivo, (2) enunciados empíricos e (3) premissas que não são nem enunciados empíricos nem regras de direito positivo”.¹⁰³⁵ Nessa linha, para a concepção de uma argumentação jurídica que leve em consideração o suporte fático e seus elementos (suas propriedades), a justificação externa pode se manifestar através de enunciados empíricos, a fim de que haja a individualização da norma jurídica. Desse modo, para se delinear o suporte fático, sempre existirá a utilização de argumentação prática do tipo geral, uma vez que para “qualquer falante e em qualquer momento é possível passar a um discurso teórico (empírico)”¹⁰³⁶, conforme ressalta Alexy.

Em síntese, falar de suporte fático no campo dos direitos é se expressar através de argumentação prática do tipo geral no campo argumentativo, impactando no modo como se pensa o desenvolvimento de sistemas de IA judicial.

4.1.2 Fatores (“*factors*”) e a Representação do Suporte Fático

Para a representação do suporte fático os sistemas de Inteligência Artificial judicial podem utilizar os **fatores** (*factors*) na forma definida no modelo desenvolvido por Trevor Bench-Capon e Giovanni Sartor¹⁰³⁷ (conhecido como *B-CS model*) ou o modelo de **conceitos jurídicos intermediados** (*intermediate legal concepts*), ou *ILC’s*, desenvolvido por Kevin Ashley.¹⁰³⁸

Primeiramente, em relação aos fatores (*factors*), estes podem representar a ocorrência de atos, fatos ou estados existentes no mundo da vida. Nesse ponto, destaca-se que o uso de fatores para a implementação de programação computacional voltada à Inteligência Artificial jurídica tem seu modelo padrão desenvolvido por Trevor Bench-Capon e Giovanni Sartor¹⁰³⁹, que ficou conhecido na comunidade acadêmica através do clássico estudo denominado *A Model of Legal*

¹⁰³⁴ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 203.

¹⁰³⁵ Id.

¹⁰³⁶ Ibid., p. 257.

¹⁰³⁷ BENCH-CAPON, Trevor; SARTOR, Giovanni. A Model of Legal Reasoning with Cases Incorporating Theories and Values.

¹⁰³⁸ ASHLEY, Kevin D. **Artificial Intelligence and Legal Analytics**, p. 73.

¹⁰³⁹ BENCH-CAPON, Trevor; SARTOR, Giovanni. Op. cit.

Reasoning with Cases Incorporating Theories and Values, ou simplesmente conhecido como o *B-CS Paper*.¹⁰⁴⁰

Originalmente, o modelo B-CS foi pensado no paradigma de desenvolvimento de raciocínio baseado em casos (*case based reasoning*), a fim de se construir teorias por meio da representação de casos e valores, para que o sistema de IA jurídica fosse capaz de decidir novos casos.¹⁰⁴¹

Desse modo, os fatores podem ser utilizados para representar **estados, fatos e atos** que, caso ocorram, levam ao acionamento da norma jurídica representada no sistema de IA judicial. No modelo B-CS, os casos passados são representados através de fatores que, por sua vez, representam razões que explicam seus resultados. Deve-se destacar que os fatores (*factors*) de Bench-Capon e Sartor se aproximam do conceito de propriedade utilizado por Bulygin e Alchourrón ao definirem o Universo das Propriedades (UP). Assim, os fatores são abstrações dos fatos que compõem os casos, o que indica que os fatores descrevem um caso.¹⁰⁴² Ou seja, no nível mais inicial do raciocínio jurídico, as ontologias e as heurísticas estão no nível dos fatores, não no nível dos conceitos jurídicos em si mesmos, ou mesmo no nível das regras.

Para facilitar a compreensão, seguem alguns exemplos de fatores:

- a) fatores que representam fatos: data, nascimento, morte, dano, etc.;
- b) fatores que representam atos: compra, venda, manifestação em um processo, afirmação, negação, etc.;
- c) fatores que representam estados: vida, intensidade do dano, lugar, tempo, modo, dúvida, etc.

Numa análise abrangente do tema, percebe-se que o acionamento da norma jurídica através de fatores (*factors*) pode ser intermediado por um caso que represente

¹⁰⁴⁰ ASHLEY, Kevin. Evaluating the Uses of Values in a Model of Legal Reasoning. In: ATKINSON, Katie; PRAKKEN, Henry; WYNER, Adam. **From Knowledge Representation to Argumentation in AI, Law and Policy Making**: A Festschrift in Honour of Trevor Bench-Capon on Occasion of his 60th Birthday. College Publications: London, 2013, p. 11.

¹⁰⁴¹ (...) “the B-CS Paper is an excellent vehicle for teaching law students a lesson that reasoning with cases is ‘a process of constructing, evaluating and applying a theory. After all, that is exactly how Bench-Capon and Sartor have modeled reasoning with cases, as constructing theories with cases and values and applying them to decide new cases” (Ibid., p. 11-12).

¹⁰⁴² “In the B-CS model, past cases are represented in terms of factors representing reasons that explain their outcomes. Factors are abstractions from the facts of the cases in the regulatory domain. The factors provide not only a way of describing cases, but the presence of a factor in a case is ‘taken to strengthen the case for in or other of the parties to the dispute.’ [8, p. 102]” (Ibid., p. 14).

sua aplicação (*case based reasoning*)¹⁰⁴³ ou pode representar o acionamento de uma regra jurídica prevista na legislação (*statutory reasoning*). Ambos os modelos, de raciocínio baseado em casos e de raciocínio baseado em regras, podem ser adotados de maneira a representar a aplicação do Direito através de IA judicial, uma vez que os paradigmas de argumentação através de casos análogos¹⁰⁴⁴ e por meio de lógica normativa¹⁰⁴⁵ integram o rol de possíveis técnicas a serem utilizadas pelo Judiciário para interpretar o Direito.

Um aspecto a ser levado em conta é que o modelo B-CS pode ser implementado de maneira a garantir poder de explicabilidade, consistência e simplicidade¹⁰⁴⁶, pois a *accountability* é estimulada pela potencialidade de compreensão do funcionamento do sistema de IA judicial.

No mais, os fatores conectam os fatos, atos ou estados do caso (propriedades ou premissas empíricas) aos valores existentes no domínio representado computacionalmente, de maneira que cada fator está associado a um ou mais valores que ele promove¹⁰⁴⁷, o que o aproxima de uma análise de precedentes.¹⁰⁴⁸

Dessa forma, a representação dos fatores em vários níveis permite a construção de sistemas de IA judicial que expliquem os resultados nos casos passados e determinem o resultado de novos casos.¹⁰⁴⁹ Assim, no modelo B-CS, um programa pode inferir casos a serem decididos através de preferências entre fatores concorrentes e entre valores concorrentes.¹⁰⁵⁰ As preferências entre os fatores são

¹⁰⁴³ (...) “to serve certain ends [and the] rules derived from cases draw their justification from the fact that following them promotes some desirable end. Thus when rules conflict, we resolve this conflict through a consideration of the purposes served and their relative desirability. Precedent decisions record the ways in which conflicts have been resolved in the past and can be seen as revealing preferences amongst different purposes. Once revealed, these preferences can be used to resolve further disputes” (BENCH-CAPON, Trevor; SARTOR, Giovanni. Op. cit.).

¹⁰⁴⁴ ALEX, Robert. **Teoria da Argumentação Jurídica**, p. 240-241.

¹⁰⁴⁵ Ibid., p. 196.

¹⁰⁴⁶ “Finally, on the B-CS model, the theories constructed can be compared in terms of their coherence, a process that the authors implement computationally in terms of explanatory power, consistency, and simplicity” (ASHLEY, Kevin. Evaluating the Uses of Values in a Model of Legal Reasoning, p. 13).

¹⁰⁴⁷ “The factors also connect the case facts to the often-competing values underlying the regulatory domain. Each factor is associated with a value that it promotes” (Ibid., p. 14).

¹⁰⁴⁸ “I will characterize this model, in which preferences between sets of values revealed in past cases, defeasibly ‘determine’ outcomes in new cases, as a ‘precedential’ model of values. Based on their repeated statements in the following quotations and elsewhere, I believe that is what Bench-Capon and Sartor intended” (Ibid., p. 14-15).

¹⁰⁴⁹ “Their multi-tiered representation enables a kind of theory construction, where the constructed theory can be used to explain the outcomes in the past cases and determine the outcome of a new case” (Ibid., p. 14).

¹⁰⁵⁰ Id.

expressas em decisões passadas, que indicam, dessa forma, as prioridades nas regras, a partir das quais podem ser inferidas preferências entre valores.¹⁰⁵¹

Todavia, muitas vezes, a abstração realizada através de fatores pode ser, de certa forma, arbitrária em relação àquilo que se observa dos elementos inerentes à construção do raciocínio baseado em casos¹⁰⁵², desde que realizadas as representações dos fatos e dos atos que normalmente acontecem através de fatores (*factors*) específicos.¹⁰⁵³ Isso demonstra que há espaço para a construção de soluções previamente delineadas no desenvolvimento de sistemas de IA judicial, o que possibilitará a atuação cooperativa humano-máquina típica do duplo canal de correção advindo do modelo de juiz ciborgue.

O uso de fatores para a programação de IA judicial pode avançar para além da análise de precedentes, pois os fatores podem representar os elementos (propriedades ou enunciados empíricos) advindos de uma visão do legítimo julgador a respeito de um caso hipotético, a fim de se alinhar previamente a uma interpretação esperada. Por exemplo, um determinado juiz pode nunca ter julgado um caso em que tenha envolvido o processamento de um *habeas data*¹⁰⁵⁴, mas pode, se assim o quiser, prever fatores (*factors*) que levariam à admissão (ou à inadmissão) do processamento da ação e ao julgamento de mérito (procedência ou improcedência).

Um exemplo mais realista ocorre quando o julgador estuda um primeiro caso e define os fatores (de improcedência ou procedência do pedido, por exemplo), e, a partir disso, programa mentalmente que na hipótese de surgir outro caso representado por determinados outros fatores, desde já haja uma decisão pré-programada. Esse tipo de abordagem passa por uma evolução quando o julgador pode transferir essa programação do espaço puramente mental para o desenvolvimento de sistemas de

¹⁰⁵¹ “*Preferences between factors are expressed in past decisions, which thus indicate priorities between these rules. From these priorities we can abduce certain preferences between values. Thus the body of case law as a whole can be seen as revealing an ordering on values*” (BENCH-CAPON, Trevor; SARTOR, Giovanni. Op. cit.).

¹⁰⁵² “*These arbitrary preferences are often required to enable a theory to justify a position when no position is determined by previous cases*” (Id.).

¹⁰⁵³ “*This assumption is not arbitrary, but rather grounded on de the evidence provided by precedent c (similar to the way in which scientific theories are grounded in the evidence provide by empirical observations). Accepting this preference between two rules also commits is to a preference for the values promoted by the preferred rule over those promoted by defeat rule. We therefore introduce a theory constrictor to include such abductions based on the evidence of previous decisions in our theories*” (Id.).

¹⁰⁵⁴ Na prática, as ações de *habeas data* são bastante raras, haja vista o manejo de outros tipos de ações que alcançam resultado análogo e que têm maior maleabilidade quanto aos pedidos que podem veiculados pelas partes do processo.

IA judicial. Observe-se que não se fala, nesse caso, em uma IA que é um simples repositório de regras e fatores que indicam casos do passado, ou mesmo de um modelo preditivo baseado em aprendizado de máquina. O que se revela é a individualização da norma jurídica para um caso hipotético, ou seja, é o Juiz que, segundo sua pré-compreensão do Direito (e sua pretensão de correção normativa e moral), molda o sistema de IA judicial. Esse tipo de abordagem pautada na definição do que fazer (como julgar) se determinadas hipóteses se realizarem é uma prática relativamente comum, pois os magistrados usualmente orientam suas assessorias como agir em determinados casos.

Kevin Ashley, por sua vez, entende que fatores (*factors*) que descrevem casos passados não necessariamente determinam valores a serem seguidos pelo sistema de IA jurídica¹⁰⁵⁵, isso porque, no modelo B-CS¹⁰⁵⁶, a descrição de todos os casos considerados relevantes no microssistema (o que envolve a identificação dos fatores) se limita a representar os casos, as regras disponíveis para serem usadas na explicação desses casos, as preferências entre as regras e os valores para serem usados na resolução de conflitos entre regras. Contudo, tudo isso foca no campo da representação detalhada dos fatores de um caso, algo que não se pode alcançar com valores *prima facie*.¹⁰⁵⁷

Nessa argumentação baseada em fatores (*factors*) defendida por Bench-Capon e Sartor, “o antecedente de uma regra é formado a partir de fatores que favorecem o resultado que forma o consequente”¹⁰⁵⁸, sendo que os fatores (*factors*) se relacionam

¹⁰⁵⁵ “It is this ‘precedential’ use of values, however, to which I object as a law professor. I do not believe that past cases ‘determine’ preferences among values or an ordering values that may be properly addressed in a reasonable legal argument to a judge deciding a new case. I believe it is conceptually misleading to characterize a teleological argument in this manner and that it overly simplifies legal arguments from values. In short, it simply is not how legal advocates and judges argue with values” (ASHLEY, Kevin. *Evaluating the Uses of Values in a Model of Legal Reasoning*, p. 15-16).

¹⁰⁵⁶ [A theory] “contains description of all cases considered relevant by the proponent of the theory, descriptions of all factors chosen to represent those cases, all rules available to be used in explaining the cases, and all preferences between rules and values available to be used in resolving conflicts between rules” (BENCH-CAPON, Trevor; SARTOR, Giovanni. Op. cit.).

¹⁰⁵⁷ “Ainda mais difícil que o problema da identificação daquilo que deve ser ordenado são os problemas relativos à ordenação em si mesma. Essa ordenação pode ser levada a cabo de duas formas: de forma cardinal ou de forma ordinal. Uma ordenação *cardinal* ocorre quando aos valores são atribuídos números a partir de uma escala, os quais expressam sua hierarquia ou seu peso. Nesse sentido, seria possível pensar em expressar a hierarquia dos valores por meio de números de uma escala que vá de 0 a 1. Já uma ordenação *ordinal* é menos exigente. Ela exige apenas que sejam estabelecidas relações de superioridade valorativa (preferência) e de igualdade valorativa (indiferença) entre os valores a serem ordenados. É fácil perceber que uma ordenação abstrata de valores constitucionais, não importa de cardinal ou ordinal, é inaceitável” (ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 159-161).

¹⁰⁵⁸ Tradução livre de: “based on factors: the antecedent of a rule is formed from factors favouring the outcome which forms consequent” (BENCH-CAPON, Trevor; SARTOR, Giovanni. Op. cit.).

com o suporte fático inerente à aplicação do Direito, bem como tem estrita ligação com os argumentos práticos gerais.

Nesse aspecto, pensar a representação do Direito através de programação de sistemas de Inteligência Artificial judicial envolve essa consideração, pois o modelo B-CS tem sua utilidade em relação à explicitação do suporte fático e dos elementos argumentativos do tipo prático geral.

Outro ponto relevante a se considerar é que os fatores (*factors*) programados podem ser a base de sugestões de aplicações do direito por analogia, o que demandará a comunicação entre os fatores de microssistemas jurídicos diferentes, de maneira que fatores que compõem a programação jurídica possam “circular” de maneira controlada entre ramos diversos do Direito, caso essa seja uma necessidade em um sistema de IA judicial específico. Obviamente, o risco (e a oportunidade) dessa livre circulação dos fatores é a maior entropia do sistema, o que demanda que a analogia seja colocada como uma das possibilidades na fase da argumentação externa.

Ademais, é possível conjugar fatores (*factors*), relacionando fatos, atos e estados de um caso hipotético com a norma proposta para decisão mediante valores. Esses fatores, por sua vez, podem ser relacionados aos conceitos jurídicos intermediados inerentes à norma jurídica representada.¹⁰⁵⁹

Destaca-se que o modelo B-CS não se baseia em hipóteses, mas numa abordagem cujo formalismo de julgamento se volta para as progressões de eventos¹⁰⁶⁰ e para representação do conhecimento relevante¹⁰⁶¹, o que passa pela representação de estados, eventos, ações e modais das proposições normativas

¹⁰⁵⁹ “It would be a theory relating the facts of a case or hypothetical to the announced or proposed rule for decision in light of the values (i.e., applicable social propositions or analogy-warranting rationale) underlying the legal rule and associated with the intermediate legal concepts” (ASHLEY, Kevin. Evaluating the Uses of Values in a Model of Legal Reasoning, p. 23).

¹⁰⁶⁰ “Before we go any further, consider that there are three distinct ways that programs can interact over a network: commands, events, and queries. If you’ve not considered the distinction between these three before, it’s well worth doing so, as it provides an important reference for interprocess communication.” (...)

“ Events are both a fact and a notification. They represent something that happened in the real world but include no expectation of any future action. They travel in only one direction and expect no response (sometimes called ‘fire and forget’), but one may be ‘synthesized’ from a subsequent event” (STOPFORD, Ben. **Designing Event-Driven Systems: Concepts and Patterns for Streaming Services with Apache Kafka**. Sebastopol: O’Reilly, 2018, p. 30-31).

¹⁰⁶¹ “The B-CS model makes no provision for posing hypotheticals. As explained below, our approach does, with its value judgment formalism and representations of event progressions and relevant common sense knowledge” (ASHLEY, Kevin. Evaluating the Uses of Values in a Model of Legal Reasoning, p. 28-29).

(obrigações, proibições, permissões e facultatividades)¹⁰⁶², decorrendo da aplicação do direito (e dos valores presentes nas normas jurídicas) em cenários específicos mediante a atribuição de consequências jurídicas para os eventos presentes nos fatores (*factors*).¹⁰⁶³

Por exemplo, imagine que uma petição não atenda a todos requisitos exigidos pelo art. 330, §1º, do Código de Processo Civil. Baseado nisso, um fator (o suporte fático) pode desencadear uma consequência jurídica (intimação do autor para emendar a petição inicial) cuja regra está pressuposta na programação (abstração e heurística), mas não está escrita na programação. Noutro exemplo, qual é o melhor meio de prova para um ato do mundo fático que deixa vestígios físicos? Um sistema de IA judicial pode ser programado para, com base na alegação de um tipo de fato pela parte, determinar uma espécie de produção probatória pela parte.

Assim, o desenvolvimento de sistemas de IA judicial leva em conta elementos de diversos paradigmas de programação, como: *domain-driven design* (baseado no estudo do domínio junto aos especialistas na área de interesse)¹⁰⁶⁴, *data-driven design* (pautado na análise de dados reais para a tomada de decisões)¹⁰⁶⁵ e *event-driven design* (considerado o suporte fático representado através de eventos que disparam a informação para processamento pelas camadas subsequentes do sistema, o que é também considerado para eventos subsequentes encadeados até um resultado).¹⁰⁶⁶ Mais do que um especialista em programação, o engenheiro jurídico (ou o assessor desenvolvedor)¹⁰⁶⁷ deve conhecer as possibilidades de cada uma dessas abordagens, pois cada problema tem seus desafios e individualidades a serem tratados.

¹⁰⁶² “I imagine that Thorne McCarty would agree with the assertion that modeling legal theory construction requires an enriched fact representation. After all, his knowledge representation for the Macomber argument, as described in Section 3 of [21, pp. 278-281], employed his Language for Legal Discourse (LLD), with its representation of states, events, actions and deontic modalities such as permission and obligation to represent the facts of the corporate distributions not only in the case itself, but in the precedents and hypotheticals” (Ibid., p. 29).

¹⁰⁶³ “The value judgement formalism [16, 17, 18] see also [2] defines and implements asset of schemes for arguing about the consequences that certain decisions have on underlying values, namely decisions to apply (or not to apply) legal rules in specific scenarios. The rules, proposed tests for deciding a case, are expressed in terms of intermediate legal concepts. The consequences are characterized in terms of whether the positive effect of a decision on the applicable legal values outweighs the negative effect. The argument schemes explicitly refer to these effects” (Id.).

¹⁰⁶⁴ EVANS, Eric. Op. cit., p. 198.

¹⁰⁶⁵ PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 15.

¹⁰⁶⁶ STOPFORD, Ben. Op. cit., p. 30-31.

¹⁰⁶⁷ PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 15.

Diante dessa inafastável conexão do direito com o que acontece no mundo, Alexy ressalta que sempre é possível retornar aos argumentos empíricos¹⁰⁶⁸, motivo pelo qual o modelo B-CS, ao se valer de fatores (*factors*) para representar eventos, estados e ações, pode ser utilizado como base para a construção de sistemas de IA judicial.

Na prática, quando se busca implementar argumentação jurídica através de sistemas de Inteligência Artificial, sempre que for necessário representar o mundo fático, os fatores (*factors*) podem ser utilizados como uma técnica possível, o que se adequa às exigências da argumentação jurídica.

Por outro lado, destaca-se que os limites da representação linguística da realidade permanecem vigentes, sendo que a consciência a respeito das limitações da representação é parte da concepção de racionalidade limitada que permeia esta obra (vide subcapítulo 3.4) e que leva em conta a natural entropia jurídica (conforme exposto no subcapítulo 1.6). Sendo assim, a partir desta abordagem, não é intenção deste estudo resolver definitivamente o problema dos limites da representação da realidade através de elementos linguísticos, haja vista que o Direito e a Inteligência Artificial não têm esse alcance.

4.1.3 Conceitos Jurídicos Intermediados (“*ILC’s*”), o Suporte Fático Hipotético e a Abordagem Híbrida

Outro modelo a ser abordado para a programação de sistema de Inteligência Artificial judicial se baseia nos **conceitos jurídicos intermediados** (*intermediate legal concepts*), que trazem para dentro da programação de sistemas de Inteligência Artificial judicial a representação da norma jurídica e de sua proposição. Ou seja, a regra jurídica é incorporada, em si mesma, na programação do sistema de Inteligência Artificial.

¹⁰⁶⁸ “Já se constatou que nos discursos práticos surgem muitos problemas que não podem ser resolvidos com os meios da argumentação prática. Pode tratar-se de questões de fato (em especial, a previsão de consequências), de problemas linguísticos (em especial, problemas de comunicação) e de questões que se referem à própria discussão prática. Nesses casos deve ser possível passar a outras formas de discurso. Esta possibilidade está assegurada mediante as seguintes regras:

(6.1) Para qualquer falante e em qualquer momento é possível passar a um discurso teórico (empírico)” (ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 182).

O modelo de *ILC's* (*intermediate legal concepts*)¹⁰⁶⁹ defendido por Kevin Ashley apresenta vantagens por possibilitar a representação heurística na argumentação jurídica (interna e externa), de maneira que o sistema de IA judicial pode ser moldado em termos de conceitos jurídicos incorporados como o conteúdo da própria programação do sistema.¹⁰⁷⁰

Desse modo, é natural que existam preferências do julgador entre as interpretações possíveis, as quais podem ser representadas em conceitos jurídicos intermediados (*ILC's*).¹⁰⁷¹ Num maior aprofundamento da questão, um caso sobre o qual foram decididas preferências a respeito de determinados valores constitucionais é a representação de uma regra baseada em fatores (*factors*), o que pode ser usado para justificar a preferência representada.¹⁰⁷² Dessa forma, é preciso mapear os argumentos dos casos passados e relacioná-los com os fatos atuais mediante uma rica representação¹⁰⁷³, de modo que se busque inferir respostas a partir disso.

Sendo assim, um sistema de IA judicial que utilize a abordagem híbrida (modelo B-CS + *ILC's*) irá se basear em esquemas de argumentos básicos (regras de senso comum, argumentos de norma e argumentos de valor), bem como empregará esquemas mais complexos, como os argumentos da *consequência desejável*, da

¹⁰⁶⁹ "The proposed rule or the intermediate legal concepts may be taken or adapted from a legal rule found in a precedent case, or an advocate may propose a new rule or invent a new concept, perhaps by analogy to some other area of law" (ASHLEY, Kevin. Evaluating the Uses of Values in a Model of Legal Reasoning, p. 15-16).

¹⁰⁷⁰ "The ILCs have content. The intermediate legal concepts embody legal values and principles, as Lindahl explained in his response to Alf Ross's skeptical view of legal concepts [20, pp. 183-4]. Legal rules fashioned in terms of intermediate legal concepts that embody underlying principles, generate transparent, normatively coherent explanations of decisions that provide guidelines for handling new and problematic cases" (Ibid., p. 17).

¹⁰⁷¹ "The preferences among rules indicate that one rule is preferred over another. The preferences are derived either from cases or from values. A decided case to which a pair of conflicting rules applies has an outcome from which one may conclude that the rule with that outcome is preferred over the other". "In addition, a decided case to which conflicting values apply (based on the values linked to factors that apply in the case) is evidence that one set of values is preferred over another, and his preference can also be used to justify a preference among factor-based rules" (Ibid., p. 26).

¹⁰⁷² "In addition, a decided case to which conflicting values apply (based on the values linked to factors that apply in the case) is evidence that one set of values is preferred over another, and his preference can also be used to justify a preference among factor-based rules" (Id.).

¹⁰⁷³ "In addition, to implement a scheme for legal argument-from-theory successfully, I suspect, a program will need a richer representation for facts than factors provide. One can use past cases as a guide to resolving conflicting values in a current case, but it is a matter of mapping arguments about the values stake in the past case to context of the current facts. One needs to compare the current facts an those of a past case I sufficient detail to make sure how well the mapped arguments apply in the new circumstances" (Ibid., p. 28).

*consequência desejável mais provável, da consequência desejável menos provável e argumentos correspondentes das consequências indesejáveis.*¹⁰⁷⁴

Dessa maneira, ocorre um passo rumo ao processo de individualização da norma jurídica através da escolha de conceitos jurídicos intermediados (ILC's) com base nos quais o legislador ou o juiz formulam uma regra jurídica, o que é uma maneira pela qual o legislativo e o judiciário tentam acomodar prospectivamente os valores constitucionais conflitantes, antecipando os cenários nos quais surgirão.¹⁰⁷⁵ Nesse sentido, as preferências entre possíveis regras surgidas do conflito de princípios podem ser representadas como parte da programação do sistema de IA judicial¹⁰⁷⁶, sem que isso impossibilite uma nova rodada de superação dos argumentos através da pretensão de correção moral e normativa advinda da razão humana no contexto do juiz ciborgue.

Nesse caminho, para além do suporte fático abstrato¹⁰⁷⁷ (da norma) e do suporte fático concreto (dos fatos), há o suporte fático hipotético vinculado àquilo que as partes buscam demonstrar através de seus discursos argumentativos. Como o suporte fático concreto leva à individualização do suporte fático abstrato, um discurso sobre o suporte fático ainda não concretizado pode ser caracterizado como suporte fático hipotético presente diretamente no sistema de IA judicial.

Essa abordagem é necessária para que as hipóteses fáticas sejam testadas e esquadrihadas por sistema de IA judicial, uma vez que, mediante a aplicação de um princípio cujo mandamento é apenas *prima facie*, consideradas determinadas condições do caso concreto (um dado suporte fático), pode surgir uma regra como mandamento definitivo.¹⁰⁷⁸ A partir disso, uma regra R' advinda de uma regra R ou de

¹⁰⁷⁴ "The program uses basic argument schemes including argument from common sense rule, argument from norm, and argument from value effect. It also employs more complex schemes such as argument from desirable consequence, from more likely desirable consequence, from less likely desirable consequence, and corresponding arguments from undesirable consequences" (Ibid., p. 31).

¹⁰⁷⁵ "The choice of intermediate legal concepts with which a legislature or judge formulates a legal rule is one way in which they attempt prospectively to accommodate the conflicting underlying social values that are anticipated in the scenarios that will arise" (Ibid., p. 17).

¹⁰⁷⁶ "Thus, ILCs are a locus of interaction between the values associated with legal principles and case facts in a regulatory domain in a way that factors are not, or have not been so far, even when factors are associated with particular values as in the B-CS model. Although the B-CS model incorporates rules based on factors, the model does not involve arguing about the meaning of those factors as intermediate legal concepts. In realistic legal arguments about cases and values, the meaning and scope of intermediate legal concepts are at issue" (Ibid., p. 18).

¹⁰⁷⁷ AFONSO DA SILVA, Virgílio. **Direitos Fundamentais**, p. 67-68.

¹⁰⁷⁸ "Decisões sobre direitos pressupõe a identificação de direitos definitivos. O caminho que vai do princípio, isto é, do direito *prima facie*, até o direito definitivo passa pela definição da relação de

um princípio P pode ser representada em termos de *conceito jurídico intermediado* em um sistema de IA judicial.

4.2 A Justificação Interna das Decisões Tomadas com Auxílio de Inteligência Artificial

A argumentação jurídica tem papel muitíssimo importante na Teoria do Direito, pois se trata de campo de estudo que no desenvolvimento jurídico se mostrou necessário a fim de atuar num paradigma de lógica racional que se espera de uma decisão judicial.¹⁰⁷⁹

Desse modo, a forma de atuação comunicacional no direito deveria ser realizada pela argumentação jurídica, que é o ônus epistemológico necessário para o mínimo de prestação de contas de como as decisões são tomadas¹⁰⁸⁰, bem como se tais decisões enfrentaram (ou simplesmente levaram em conta) os argumentos das partes do processo (agentes de comunicação¹⁰⁸¹).

A linguagem natural (português, inglês, espanhol ou outro idioma), falada ou escrita, é o meio de comunicação padrão que embasa o grande sistema operacional do Direito. Argumentar é, sobretudo, se comunicar. Conforme define Habermas em sua Teoria do Agir Comunicativo, a linguagem se demonstra como processo, procedimento e como lógica de justificação.¹⁰⁸²

Nessa linha, apesar da existência concomitante de visões bastante otimistas¹⁰⁸³ e pessimistas¹⁰⁸⁴ acerca dos avanços da tecnologia da Inteligência Artificial e sua aplicabilidade ao Direito, o surgimento de técnicas de Inteligência Artificial com capacidade de propositura de soluções para diferentes tipos de casos,

preferência. Mas a definição de uma preferência é, segundo a lei da colisão, a definição de uma regra” (ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 107).

¹⁰⁷⁹ ATIENZA, Manuel. Op. cit., p. 254-258.

¹⁰⁸⁰ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 36-37.

¹⁰⁸¹ HABERMAS, Jürgen. **Teoria do Agir Comunicativo**, 1, p. 48.

¹⁰⁸² Ibid., p. 60-62.

¹⁰⁸³ CELLA, José Renato Gaziero. Op. cit., p. 137.

¹⁰⁸⁴ “Por que não confiar todas as decisões difíceis a máquinas que sejam capazes de processar por nós quantidades imensas de dados? Podemos agir assim por decisão consciente, desejando uma vida mais fácil; ou podemos fazê-lo inconscientemente, porque nossa dependência cada vez maior das máquinas nos tornou incapazes de saber quando parar. Todos conhecemos os sinais. Passamos horas enviando e respondendo e-mails insignificantes não em obediência aos nossos computadores, mas simplesmente porque nos falta a capacidade de quebrar o encanto. A única coisa que nos liberta é o surgimento de uma tecnologia ainda mais acessível e imediata. Aí nos viciamos nela. A conveniência sem custo é sua própria maldição” (RUNCIMAN, David. Op. cit., p. 133-134).

inclusive para a tomada de decisões judiciais¹⁰⁸⁵, é uma realidade. É verdade que muito deve ser melhorado e desenvolvido, mas o grau de acurácia de decisões propostas por meio de Inteligência Artificial gera debate a ponto de haver resistência de parcela dos juristas ao mostrar desalento com o novo.

A Inteligência Artificial no Direito é uma realidade já existente no Supremo Tribunal Federal¹⁰⁸⁶, no Superior Tribunal de Justiça¹⁰⁸⁷, nos Tribunais de Justiça de Minas Gerais¹⁰⁸⁸ e de Rondônia¹⁰⁸⁹, com processo de implantação no Tribunal de Justiça do Estado do Paraná¹⁰⁹⁰. O Conselho Nacional de Justiça¹⁰⁹¹ passou a estimular a adoção de Inteligência Artificial no Poder Judiciário.

Para buscar lançar as linhas do debate, o presente tópico se pauta na necessidade de uma interface entre teoria da argumentação jurídica e os sistemas de Inteligência Artificial. Como recorte teórico necessário, o presente tópico traz o modelo de teoria da argumentação jurídica de Alexy, a fim de que haja uma possibilidade de conversação multidisciplinar entre Teoria do Direito e Ciência da Computação. Neste aspecto, a possibilidade de utilização da teoria da argumentação jurídica de Alexy para a decisão judicial mediante Inteligência Artificial é discutida no meio científico¹⁰⁹²,

¹⁰⁸⁵ MINAS GERAIS. Op. cit.

¹⁰⁸⁶ BRASIL. Supremo Tribunal Federal. **Inteligência artificial vai agilizar a tramitação de processos no STF**. [2018]. Disponível em: <http://www.stf.jus.br/portal/cms/verNoticiaDetalhe.asp?idConteudo=380038>. Acesso em: 01 set. 2019.

¹⁰⁸⁷ BRASIL, Superior Tribunal de Justiça. **STJ entra na era da inteligência artificial**. [2018]. Disponível em: http://www.stj.jus.br/sites/portallp/Paginas/Comunicacao/Noticias-antigas/2018/2018-06-14_06-51_STJ-entra-na-era-da-inteligencia-artificial.aspx. Acesso em: 01 set. 2019.

¹⁰⁸⁸ MINAS GERAIS. Op. cit.

¹⁰⁸⁹ RONDÔNIA. Tribunal de Justiça do Estado de Rondônia. **Inteligência Artificial desenvolvida pelo TJRO pode revolucionar o Judiciário**. [2018]. Disponível em: <https://www.tjro.jus.br/noticias/item/9472-inteligencia-artificial-desenvolvida-pelo-tjro-pode-revolucionar-o-judiciario>. Acesso em: 01 set. 2019.

¹⁰⁹⁰ PARANÁ. Tribunal de Justiça do Estado do Paraná. **TJPR vai lançar novos sistemas de inteligência artificial**. [2019]. Disponível em: <https://www.tjpr.jus.br/noticias>. Acesso em: 01 set. 2019.

¹⁰⁹¹ CONSELHO NACIONAL DE JUSTIÇA. **CNJ anuncia a criação de laboratório de inteligência artificial para o PJe**. [2018]. Disponível em: <https://www.cnj.jus.br/noticias/cnj/88132-cnj-anuncia-a-criacao-de-laboratorio-de-inteligencia-artificial-para-o-pje>. Acesso em: 31 ago. 2019.

¹⁰⁹² “Parece cristalino o fato de que Alexy delega aos sistemas jurídicos inteligentes essa possibilidade [...]. Quando ocorre de a decisão de um caso singular não se seguir logicamente nem das normas pressupostas nem de enunciados solidamente fundamentados de um sistema qualquer - justamente com enunciados empíricos -, nem poder ser fundamentada definitivamente com a ajuda das regras da metodologia jurídicas, então, resta ao intérprete um campo de ação senão terá de escolher entre várias soluções – pela inteligência artificial, a partir de normas jurídicas, regras metodológicas e enunciados de sistemas jurídicos inteligentes será possível ponderar”. [...] “Bem, alguns teóricos de inteligência artificial e de Direito têm tomado a distinção de justiça feita anteriormente por Rawls para afirmar a natureza processual do argumento jurídico” (MOZEIC, Vinícius Almada. Os sistemas jurídicos inteligentes e o caminho perigoso até a teoria da argumentação de Robert Alexy. **Revista Brasileira de Direito**, Passo Fundo, v. 13, n. 3, 2017).

ainda que seja combatida sem aprofundamento nas possibilidades do efetivo entrelaçamento entre a argumentação jurídica e a Ciência da Computação.

Alexy divide a argumentação jurídica nos níveis de argumentação interna e argumentação externa.¹⁰⁹³ Este tópico se preocupará com o primeiro nível referido, haja vista ser o ponto de partida que o intérprete do direito deve respeitar no processo argumentativo.

A teoria argumentação de Alexy tem pensamento procedimentalista, razão pela qual percebe a interpretação do direito como um processo racional de justificação de um caso especial de proposições normativas (possibilidades de interpretação normativa) que consiste na decisão jurídica.¹⁰⁹⁴ Para tanto Alexy adota um rol de técnicas de justificação as quais enumera de J.1.1 a J.18.¹⁰⁹⁵ Todavia, no presente tópico, há um enfoque nas técnicas de justificação interna (J.1.1 a J.2.5), haja vista ser o modelo de justificação padrão que qualquer sistema que se preze ser chamado de juridicamente “inteligente” deve vencer antes de avançar à utilização de argumentos de justificação externa (dentro da argumentação jurídica) ou de argumentos práticos gerais.

Na justificação interna, “verifica-se se a decisão se segue logicamente das premissas que se expõe como fundamentação”¹⁰⁹⁶, motivo pelo qual o primeiro método a ser tentado é o do silogismo jurídico, o qual Alexy identifica como (J.1.1).¹⁰⁹⁷

Nesse caso, o silogismo (J.1.1) satisfaz o princípio da universalidade (1.3') das regras fundamentais do discurso prático geral, segundo o qual “(1.3') Todo falante só pode afirmar os juízos de valor e de dever que afirmaria dessa mesma forma em todas

¹⁰⁹³ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 195.

¹⁰⁹⁴ Ibid., p. 188-194.

¹⁰⁹⁵ Ibid., p. 257-260.

¹⁰⁹⁶ Ibid., p. 218.

¹⁰⁹⁷ Adaptando-se o exemplo dado por Alexy:

(J.1.1) .(1) (x) (x→ORx)
 .(2) Ta
 .(3) ORa (1),(2)

X: variável de indivíduo no domínio das pessoas naturais ou jurídicas.

A: é uma constante de indivíduo (p.e.: um nome próprio).

T: é um predicado que representa a suposição de fato da norma (1) enquanto propriedade de pessoas.

R: é predicado que expressa que o que tem de fazer o destinatário da norma.

Por exemplo:

(1) O soldado deve dizer a verdade em questões de serviço (§13, ap. I, da Lei do Soldado)

(2) O senhor M é um soldado

(3) O senhor M deve dizer a verdade em questões de serviço, (1) (2).

(Ibid., p. 196).

as situações em que afirme que são iguais em todos os aspectos relevantes”¹⁰⁹⁸. Isso leva à observância do princípio de justiça formal, que obriga a tratar da mesma maneira todos os que estão numa mesma categoria.

Desse modo, o segundo passo no processo argumentativo se volta à concreção do princípio da universalidade, na medida em que “Para a fundamentação de uma decisão jurídica deve-se apresentar pelo menos uma norma universal” **(J.2.1)**¹⁰⁹⁹ e que “a decisão jurídica deve seguir-se logicamente ao menos de uma norma universal, junto a outras proposições” **(J.2.2)**¹¹⁰⁰. Nesse sentido, as decisões tomadas com auxílio de Inteligência Artificial devem guardar estrita pertinência com o princípio da universalidade, que é a condição de racionalidade¹¹⁰¹ e de mínima justiça no processo de justificação das decisões judiciais.

De todo modo, Alexy destaca que (J.2.1) e (J.2.2) não devem ser superestimadas, pois não excluem a possibilidade da existência de uma cláusula de exceção, a qual também deve valer universalmente na medida em que faz parte da lógica da própria regra excepcionalizada.¹¹⁰² Dessa forma, (J.2.1) e (J.2.2) valem para os casos em que existe norma de direito positivo e “também nos casos em que não haja tal norma de direito positivo”¹¹⁰³ (“Se não se pode extrair nenhuma regra da lei, então se deve construí-la”¹¹⁰⁴). Por isso, dentro dos estreitos limites da obra, não será abordada a hipótese de construção de solução de justificação interna nos casos em que se demanda construção em decorrência de inexistência de norma jurídica advinda de fonte legislativa primária.

Nessa linha, tem-se ciência de que o esquema de fundamentação (J.1.1) é insuficiente nos “casos complicados”¹¹⁰⁵ a) quando a norma contém diversas propriedades alternativas do fato hipotético; b) quando sua aplicação exige um complemento através de normas jurídicas explicativas, limitativas ou extensivas; c) quando são possíveis diversas consequências jurídicas; e d) quando na formulação

¹⁰⁹⁸ Ibid., p. 255.

¹⁰⁹⁹ Ibid., p. 196.

¹¹⁰⁰ Ibid., p. 197.

¹¹⁰¹ “Esse modelo de Direito em três níveis (o das regras, o dos princípios e o dos procedimentos) não permite alcançar sempre uma única resposta correta para cada caso, mas é também o que leva a um maior grau de racionalidade incorporado no Direito moderno e, em particular, no Direito de um Estado democrático e constitucional” (ATIENZA, Manuel. Op. cit., p. 218).

¹¹⁰² ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 197.

¹¹⁰³ Id.

¹¹⁰⁴ Id.

¹¹⁰⁵ ATIENZA, Manuel. Op. cit., p. 205.

da norma são utilizadas expressões que admitem diversas interpretações.¹¹⁰⁶ É exatamente em tal ponto que há o primeiro grande desafio para o que se pretende denominar como Inteligência Artificial jurídica, tendo em vista que nos casos complicados está o cerne da problematização da argumentação para discurso de justificação interna, o que é inerente à própria norma jurídica (categorizada, por exemplo, em regras e princípios¹¹⁰⁷) que possui uma multiplicidade de variáveis.¹¹⁰⁸

Inspirando-se no exemplo dado por Alexy (porém o adaptando à realidade da legislação brasileira), pode ser citado como *caso complicado* a seguinte exemplificação. No caso, A apresentou exceção de pré-executividade¹¹⁰⁹ com alegação de matéria já analisada na fase de conhecimento do processo. O julgador tem de provar a aplicação da norma que prevê que “o juiz condenará o litigante de má-fé a pagar multa ...” (art. 81 do Código de Processo Civil).

A norma, segundo aplicação da teoria de Alexy¹¹¹⁰, formula-se pela seguinte lógica: **(1) (x) (Tx → ORx)**.

T está definido no art. 80 do CPC¹¹¹¹ por meio de sete características (**M¹₁-M¹₇**).

A representação do art. 80 e do art. 81 do CPC pode ser reduzida à seguinte forma: **(2) (x) (M¹₁x v M¹₂x v ... M¹₇x ↔ Tx)**.

De (1) e (2) se segue: **(3) (x) (M¹₁x v M¹₂x v ... M¹₇x → ORx)**.

(3) Se acontece alguma das propriedades se produz a consequência jurídica ORx. M¹₄ significa: “opuser resistência injustificada ao andamento do processo”.

¹¹⁰⁶ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 197.

¹¹⁰⁷ ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 90.

¹¹⁰⁸ CARNEIRO, Claudio; LEMOS, Marcos Antonio Fernandes. O Movimento Pós-Positivista e a “Visão” Neoconstitucionalista da Ponderação: A Valoração de Princípios em Robert Alexy. **Revista de Direitos Fundamentais & Democracia**, Curitiba, v. 12, n. 12, p. 111, jul./dez. 2012.

¹¹⁰⁹ A exceção de pré-executividade ainda é instrumento de defesa do executado bastante utilizado no rito da execução fiscal e no âmbito dos Juizados Especiais Cíveis, na medida em que em ambos os casos inexistente previsão específica de impugnação à execução (nos moldes do art. 525 do CPC) e os embargos à execução ocorrem somente após a garantia do juízo (art. 16, §1º, da Lei 6.830/1980 e art. 53, §1º, da Lei 9.099/1995).

¹¹¹⁰ ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 220.

¹¹¹¹ Código de Processo Civil:

“Art. 80. Considera-se litigante de má-fé aquele que:

I - deduzir pretensão ou defesa contra texto expresso de lei ou fato incontroverso;

II - alterar a verdade dos fatos;

III - usar do processo para conseguir objetivo ilegal;

IV - opuser resistência injustificada ao andamento do processo;

V - proceder de modo temerário em qualquer incidente ou ato do processo;

VI - provocar incidente manifestamente infundado;

VII - interpor recurso com intuito manifestamente protelatório”.

(BRASIL. **Lei 13.105, de 16 de março de 2015**).

Assim, tendo em vista a subsunção ao inciso IV do art. 80 do CPC no caso, de (3) se segue agora: **(4) (x) ($M^1_4x \rightarrow ORx$)**.

Segundo a definição usada pelos tribunais, opõe resistência injustificada ao andamento do processo quem *apresenta manifestações cujos fundamentos já foram anteriormente analisados por decisão definitiva que leve à preclusão*¹¹¹² **(M^2_4)**. Desse modo, a hipótese legal do art. 80, IV do CPC, aplicada ao caso **(M^1_4x)**, se decompõe numa segunda significação específica representada pela interpretação do STJ no exemplo¹¹¹³ **(M^2_4)**.

Portanto: **(5) (x) ($M^2_4x \leftrightarrow M^1_4x$)**. Ou seja, as hipóteses de aplicação são equivalentes.

A hipótese delineada pelo Superior Tribunal de Justiça no exemplo (M^2_4), por seu lado, acontece se a parte *apresenta exceção de pré-executividade que reitera matéria já julgada em definitivo* **(M^3_4)**. Portanto vale: **(6) (x) ($M^3_4x \leftrightarrow M^2_4x$)**. Mais uma vez as hipóteses de aplicação são equivalentes.

Assim, se alguém apresentar uma exceção de pré-executividade que proponha o julgamento de matéria preclusa (S), deve ser considerado como alguém que **cumpr**e M^3_4 : **(7) (x) ($Sx \rightarrow M^3_4x$)**, ou seja, mediante uma ligação retrospectiva entre M^3_4 , M^2_4 , M^1_4 e Tx , o caso S enseja a tipificação da litigância de má-fé (Tx).

Em suma, estabelecido o pressuposto: **(8) Sa**.

De (1)-(7) se segue: **(9) ORa**.

¹¹¹² Trecho do voto do Ministro Relator do AgInt no REsp 1786179/SP:

“O magistrado não poderia se quedar inerte à reiterada recalcitrância da parte em cumprir a coisa julgada, formulando pretensão destituída de fundamentos, sobretudo quando a norma processual em vigor oferece instrumentos para fazer cumprir decisões judiciais e punir os litigantes de má-fé, como no caso presente” (BRASIL. Superior Tribunal de Justiça (2. Turma), **Agravo Interno no Recurso Especial 1786179/SP**. Não cabimento de Recurso Especial em face de acórdão de ação rescisória cujo fundamento de julgamento se baseou na aplicação de dispositivo constitucional. Brasília, DF: Superior Tribunal de Justiça, [2019]. Recorrente: Zahir Dornaika e outros. Recorrido: Fazenda do Estado de São Paulo. Relator: Min. Herman Benjamin, 06 de junho de 2019. Disponível em: https://scon.stj.jus.br/SCON/GetInteiroTeorDoAcordao?num_registro=201803254544&dt_publicacao=18/06/2019. Acesso em: 14 nov. 2021).

¹¹¹³ E é apenas um exemplo, poderia ser usada uma definição de outro tribunal, ou mesmo uma definição doutrinária, o que denota que a justificação interna não é lastreada apenas em aspectos silogísticos.

Destaca-se que as premissas utilizadas na fundamentação de (9)¹¹¹⁴ ainda podem sofrer correção na fase de justificação externa¹¹¹⁵, aspecto da teoria da argumentação que não é, neste momento, inteiramente abordado, haja vista o caráter parcial de representação que é inerente à justificação interna.

Dessa maneira, observa-se que o raciocínio de justificação interna é um processo de concreção progressiva da norma, conforme se vê:

(3) (x) (M11x v M12x v ... M17x → ORx)	(de [1] e [2])
(4) (x) (M14x → ORx)	(de [3])
(5') (x) (M24x → M14x)	(de [3] e [5])
(6') (x) (M34x → M24x)	(de [5'] e [6])
(7') (x) (Sx → M34x)	(de [6] e [7])

Nesse sentido, no caso, as fases de desenvolvimento da lógica (5), (6) e (7) podem ser entendidas como regras para uso das expressões utilizadas nas etapas da fundamentação¹¹¹⁶, uma vez que as mesmas não são expressas no texto legal. A esse respeito, as regras de uso das palavras podem se classificar em forma forte e forma fraca¹¹¹⁷ da seguinte maneira: a) forma forte: **(x) (Fx ↔ Gx)**¹¹¹⁸; b) forma fraca: **(x) (Fx → Gx)**.¹¹¹⁹

É importante ser ressaltado que as premissas inerentes à forma forte podem ser substituídas pelas premissas de forma fraca, motivo pelo qual a forma fraca pode ser vista como a forma padrão a ser utilizada.¹¹²⁰ Assim, as premissas de forma fraca são a forma padrão na medida em que são o modelo mais comumente encontrado, pois nelas, embora o conceito baseado no uso da linguagem para proposição

¹¹¹⁴ A compilação do desenvolvimento fica assim representada:

(1) (x) (Tx → ORx)
 (2) (x) (M¹_{1x} v M¹_{2x} v ... M¹_{7x} ↔ Tx)
 (3) (x) (M¹_{1x} v M¹_{2x} v ... M¹_{7x} → ORx)
 (4) (x) (M¹_{4x} → ORx)
 (5) (x) (M²_{4x} ↔ M¹_{4x})
 (6) (x) (M³_{4x} ↔ M²_{4x})
 (7) (x) (Sx → M³_{4x})
 (8) Sa
 (9) ORa

¹¹¹⁵ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 199.

¹¹¹⁶ Ibid., p. 200.

¹¹¹⁷ Id.

¹¹¹⁸ Fx é equivalente a Gx, e vice-versa.

¹¹¹⁹ Fx tem como consequência Gx.

¹¹²⁰ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 200.

normativa seja naturalmente dela decomposto, não há exigência de que do uso linguístico prático (conceito) seja extraída logicamente a premissa (proposição normativa).

A respeito dos aspectos linguísticos envolvidos deve-se observar que se não for certo que o fato (A^{1121}) tenha equivalência com a previsão normativa (T^{1122}), não é aplicável a forma simples de justificação (J.1.1), devendo apresentar regras da forma padrão para o uso de T^{1123} , mediante o desdobrando de variáveis linguísticas dos conceitos utilizados na construção gramatical.

Sem a observância das regras de uso da forma padrão (variabilidade linguística das hipóteses de aplicação da norma) indivíduos A e B , iguais em todos os sentidos, podem ser tratados às vezes como T e às vezes como $\text{não-}T^{1124}$, o que contradiz o princípio da universalidade (1.3')¹¹²⁵ inerente à concepção de justiça decorrente da isonomia de tratamento perante as instituições estatais¹¹²⁶, vetor interpretativo que deve ser garantido no sistema decisório brasileiro em respeito ao art. 5º, *caput*, da Constituição Federal (“Todos são iguais perante a lei...”).¹¹²⁷

Logo, a exigência de se apresentar as regras de uso das palavras pode se basear no princípio da universalidade: quando é duvidoso se $\underline{A}$ é um $\underline{T}$, bem como quando existem dúvidas em outros níveis da fundamentação¹¹²⁸. Assim, é necessária a terceira justificativa interna que consiste em “Sempre que houver dúvida sobre se A é um T ou M^1 , deve-se apresentar uma regra que decida a questão” (**J.2.3**). Isso pode ser alcançado por meio de algoritmos de processamento de linguagem natural¹¹²⁹ (e

¹¹²¹ A apresentação de exceção de pré-executividade com a alegação de matéria preclusa, conforme exemplo utilizado.

¹¹²² No exemplo têm-se as hipóteses legais de caracterização de litigância de má-fé previstas no art. 80 e no art. 81 do CPC.

¹¹²³ ALEX, Robert. **Teoria da Argumentação Jurídica**, p. 200.

¹¹²⁴ Id.

¹¹²⁵ “(1.3') Todo falante só pode afirmar os juízos de valor e de dever que afirmaria dessa mesma forma em todas as situações em que afirma que são iguais em todos os aspectos relevantes” (Ibid., p. 193)

¹¹²⁶ RAWLS, John. **A Theory of Justice**, p. 55.

¹¹²⁷ Aqui se evoca a ideia de igualdade formal perante a lei. Por outro lado, a concepção de justiça pautada em igualdade material depende, para sua aplicação, notadamente de argumentos práticos do tipo geral voltados à análise dos critérios de discriminação de pessoas ou situações (ÁVILA, Humberto. Op. cit., p. 137-138), motivo pelo qual a igualdade material não é aplicada na fase da justificação interna. Desse modo, o estudo da igualdade material como elemento jurídico ativo na utilização de sistemas de inteligência artificial judicial depende aspectos não necessariamente relacionados à justificação interna.

¹¹²⁸ ALEX, Robert. **Teoria da Argumentação Jurídica**, p. 200.

¹¹²⁹ CHALKIDIS, Ilias; KAMPAS, Dimitrios. Deep Learning in Law: Early Adaptation and Legal Word Embeddings Trained On Large Corpora. **Artificial Intelligence and Law**, Dordrecht, v. 27, p. 171-198, jun. 2019. DOI: 10.1007/s10506-018-9238-9.

outros sistemas de IA), que é tecnologia com relevante desenvolvimento na atualidade e com promissor horizonte no Direito.

Além disso, o princípio da universalidade ganha contornos próprios dentro da argumentação jurídica, na medida em que “A decisão jurídica deve seguir-se logicamente ao menos de uma norma universal, junto a outras proposições” **(J.2.2)**.¹¹³⁰ Nesse aspecto, os sistemas computacionais de decisão judicial com auxílio de Inteligência Artificial devem ter capacidade de processamento linguístico da(s) norma(s) jurídica(s) em aplicação, uma vez que o argumento normativo é essencial para a configuração da justificação interna como uma fase mínima a evitar abusos do intérprete *ex machina* em ofensa ao princípio da reserva legal (art. 5º, II, da Constituição Federal), haja vista que aplicar a lei¹¹³¹ é um ato democrático, o que passa pela legitimidade do alcance das decisões legislativas à vida em sociedade (art. 1º, *caput*, e art. 2º da CF). Desse modo, conforme destaca Frederick Schauer, o aspecto democrático da aplicação de regras se pauta numa concepção de justiça¹¹³², pois repele o particularismo exacerbado ao garantir que casos substancialmente semelhantes tenham tratamento semelhante.¹¹³³ A aplicação de (J.2.3) juntamente com (J.2.2) leva a um tipo básico de justificação interna (J.1.2)¹¹³⁴ que pode ser implementado por meio de técnicas de Inteligência Artificial.¹¹³⁵

A proposição lógica acima está representada de maneira simplificada porque não é levada em conta a possibilidade de estruturas mais complicadas de hipóteses de fato e consequências jurídicas, o que demonstra a estrutura lógica dos casos em que a descrição específica dos fatos (Sa) é inserida no interior de um conceito geral por meio de diversos passos.

¹¹³⁰ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 258.

¹¹³¹ A aplicação da lei é aqui referida como um caminho a ser seguido pelos sistemas de inteligência artificial judicial, a fim de que exista um caminho inicial (e seguro) para as proposições normativas realizadas. Dessa forma, restam preservada a hipótese de controle de constitucionalidade ou de adaptabilidade interpretativa inerentes à legitimidade judicial democrática.

¹¹³² SCHAUER, Frederick. Op. cit., p. 135.

¹¹³³ Ibid., p. 136-137.

¹¹³⁴ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 200.

¹¹³⁵ (J.1.2)

.(1) (x) (Tx → ORx)

.(2) (x) (M¹x → Tx)

.(3) (x) (M²x → (M¹x))

[...]

.(4) (x) (Sx → Mⁿx)

.(5) As

.(6) ORa (1)-(5)

Torna-se relevante a questão, especialmente para evitar ou minorar o risco de explosão combinatorial¹¹³⁶, a respeito de quantas etapas de desenvolvimento são necessárias para fins de composição dos conceitos linguísticos. Nesse sentido, Alexy destaca que “São necessárias as etapas de desenvolvimento que permitam formular expressões cuja aplicação ao caso em questão não seja discutível” (J.2.4).¹¹³⁷ Assim, M^n se converte em S . Com poucas etapas de desenvolvimento demonstra-se a precisão do conteúdo normativo, mas esses são facilmente atacáveis (tais ataques são frequentemente não específicos). Já a realização de muitas etapas de desenvolvimento leva ao esclarecimento progressivo do caso em análise, mas, como alerta Alexy, causa maior complexidade e lentidão no processo de decisão.¹¹³⁸ Ademais, Alexy indica que “Deve-se articular o maior número possível de etapas de desenvolvimento” (J.2.5).¹¹³⁹

O referido pensamento, com as devidas adequações, pode ser transportado para a fase de decomposição de fatores variáveis da justificação interna, pois os sistemas de Inteligência Artificial devem buscar a capacidade de se trabalhar com o maior número de variáveis possíveis. Um intérprete humano ineficiente (que não tenha conhecimento e técnica adequados¹¹⁴⁰) não é parâmetro a ser balizado com os sistemas computacionais, pois uma IA jurídica deve ter parâmetros que levem a um grau de acurácia elevado o suficiente para se ter segurança acerca a interpretação

¹¹³⁶ “Para superar a explosão combinatorial, são necessários algoritmos que explorem a estrutura no domínio alvo e tirem vantagem de conhecimento prévio usando busca heurística, planejamento e representações abstratas flexíveis — capacidades que ainda estavam pouco desenvolvidas nos sistemas iniciais de IA” (BOSTROM, Nick. Op. cit., posição 456).

¹¹³⁷ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 201.

¹¹³⁸ Id.

¹¹³⁹ Id.

¹¹⁴⁰ A teoria do direito possui um rol mínimo de atributos do direito que sempre devem ser observados (fontes, normas, interpretação e justificação). A interpretação do direito é campo vasto onde existem pontos de desentendimento doutrinário e de disputas teóricas, muitas vezes, apaixonadas, mas isso não é um passe livre para a ausência de justificação ou para uma leitura convenientemente tópica de um sistema jurídico que é guiado pelo respeito ao conteúdo normativo e à coerência. O debate aqui pode ser estendido para os contrapontos entre [a] positivismo e pós-positivismo, [b] objetivistas e subjetivistas; [c] interpretação como descoberta de sentido e como atribuição de sentido; [d] construtivismo e desconstrutivismo; [e] as correntes que entendem que texto e norma se distinguem e os que pensam o contrário; [f] republicanistas, comunitaristas e liberais; dentre muitos outros temas jurídicos e filosóficos. Uma nota de rodapé não resolve tais disputas teóricas e não apaga a importância de cada uma das visões, por vezes, antagônicas. Os estudiosos do Direito não resolveram suas disputas. Não será a inteligência artificial que o fará. Os pés no chão e o sentimento instigante da pesquisa movem o presente estudo, ciente de que a perfeição não será alcançada, assim como nenhum dos grandes juristas do presente e do passado a alcançou. O presente estudo sonda as frestas do tempo, buscando explicar o fenômeno da IA e suas possibilidades, sabendo-se que há muito a ser debatido e desenvolvido.

proposta e da possibilidade de adoção da mesma por um supervisor¹¹⁴¹ humano¹¹⁴². Desse modo, merece destaque que o presente tópico não toma como base a possibilidade de uma Inteligência Artificial do tipo geral¹¹⁴³, mas sim uma Inteligência Artificial especializada em interpretação jurídica com ônus de “justificação ostensiva”.¹¹⁴⁴

Nesse ponto, a decomposição dos conceitos linguísticos por meio de processamento de linguagem natural¹¹⁴⁵ é um desafio a ser enfrentado. A esse respeito Kevin Ashley aponta que tal problema pode ser enfrentado por meio de modelos de *machine learning* voltados à identificação das características linguísticas do texto legal e de sua contextualização no uso pelos intérpretes humanos, possibilitando predição e classificação dos sentidos e usos possíveis das palavras utilizadas no texto¹¹⁴⁶, os quais são necessários à tomada de decisão. Tal problema não é tratado, neste momento, sob o aspecto das técnicas de *machine learning* do ponto de vista da Ciência da Computação, haja vista o escopo notadamente jurídico da presente obra.

Todavia, nos limites da tecnologia da Inteligência Artificial, deve ser destacado o “uso não-justificativo”¹¹⁴⁷ dos argumentos dogmáticos, que podem ser utilizados mesmo que eles próprios não sejam também fundamentados, que é o que ocorre “quando o enunciado dogmático não é posto em dúvida”¹¹⁴⁸. Dessa forma, quando a aplicação de uma regra demandar baixo nível de complexidade a justificação de sua utilização pode ser dispensada na argumentação. Aqui, mais uma vez trazendo a

¹¹⁴¹ HARTMANN PEIXOTO, Fabiano; SILVA, Roberta Zumblick Martins da. Op. cit., p. 37-39.

¹¹⁴² A supervisão humana é condição jurídica, ética e técnica de desenvolvimento do sistema de inteligência artificial jurídica, pois no atual estágio tecnológico são necessárias etapas éticas nos processos de verificação, validação, segurança e controle, sendo que neste último caso deve existir análise humana acerca dos resultados.

¹¹⁴³ “Máquinas com inteligência geral comparável à dos humanos — ou seja, dotadas de bom senso e capacidade real de aprender, raciocinar e planejar a superação de desafios complexos de processamento de informação em uma vasta gama de domínios naturais e abstratos — têm sido esperadas desde a invenção dos computadores, na década de 1940” (BOSTROM, Nick. Op. cit., posição 355).

¹¹⁴⁴ MACCORMICK, Neil. **Argumentação Jurídica e Teoria do Direito**, p.19.

¹¹⁴⁵ HARTMANN PEIXOTO, Fabiano; SILVA, Roberta Zumblick Martins da Silva. Op. cit. 9, p. 82-84.

¹¹⁴⁶ “Applying ML to legal texts will play key roles in a legal app for cognitive computing. One goal in cognitive computing is for ML algorithms to learn to identify patterns of textual features that are important for human problem-solving. This also involves identifying contexts and problems for which these patterns are most useful and the strengths of association between the observed and outcome features. The cognitive computing environment can then use these patterns for prediction and classification as human users solve new problems” (ASHLEY, Kevin. **Artificial Intelligence and Legal Analytics**, p. 214).

¹¹⁴⁷ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 228.

¹¹⁴⁸ Id.

compreensão de Alexy ao comentar a justificação externa, até certo ponto, o uso não-justificativo é obrigatório, pois “Nenhum dogmático pode fundamentar todos os enunciados dogmáticos em que se apoia para a discussão do seu problema”¹¹⁴⁹, uma vez que ninguém pode fundamentar tudo simultaneamente.

Desse modo, devem existir algoritmos de análise da complexidade dos conceitos do texto legal, conhecidos como algoritmos de processamento de linguagem natural (*natural language processing* ou *NLP*)¹¹⁵⁰, para definir o grau de fechamento/abertura de cada conceito inerente às palavras utilizadas.

Dessa maneira, a acurácia não passa apenas pelo resultado final da interpretação (decorrente da descoberta de um sentido), mas pela existência de uma justificação ostensiva¹¹⁵¹, com nível de persuasão¹¹⁵² e com sustentação lógica minimante baseados em razão prática.¹¹⁵³ Uma IA que tenha baixo nível de acurácia em suas propostas de justificação é uma IA à margem do direito, do Código de Processo Civil e da Constituição. No Direito, a ausência de adequada argumentação na justificação das decisões judiciais é inconstitucional (art. 93, IX, da CF). Por melhor que seja a justificação apresentada pela Inteligência Artificial caberá à supervisão humana¹¹⁵⁴ (juiz e sua assessoria especializada) a validação do resultado da interpretação.¹¹⁵⁵

Nesse sentido, o aspecto democrático da justificação das decisões judiciais necessita do elemento humano¹¹⁵⁶, pois é inviável a simples automatização das decisões (sem possibilidade de participação ou influência humanas), na medida em que uma “máquina distributiva que só precisaríamos programar corretamente”¹¹⁵⁷ levaria ao fato de que a justiça não seria compreendida como uma realização humana através dos agentes envolvidos, o que transformaria esses em meros beneficiários

¹¹⁴⁹ Id.

¹¹⁵⁰ HARTMANN PEIXOTO, Fabiano; SILVA, Roberta Zumblick Martins da. Op. cit., p. 33.

¹¹⁵¹ MACCORMICK, Neil. **Argumentação Jurídica e Teoria do Direito**, p.127.

¹¹⁵² ATIENZA, Manuel. Op. cit., p. 257.

¹¹⁵³ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 251.

¹¹⁵⁴ DEWEY, Daniel; RUSSEL, Stuart; TEGMARK, Max. Op. cit.

¹¹⁵⁵ HARTMANN PEIXOTO, Fabiano; SILVA, Roberta Zumblick Martins da. Op. cit., p. 39.

¹¹⁵⁶ “Pois, se consideramos as pessoas sujeitos ao mesmo tempo sociais e autônomos, que podem (ou seja, deveriam ser capazes de) determinar ativamente as estruturas normativas que devem valer para elas, então este é um padrão ao mesmo tempo imanente e transcendente à práxis: ninguém esgota completamente a práxis de justificação na qual ele ou ela participou, pois sempre há a possibilidade de questioná-la e criticá-la de modo reflexivo. Uma crítica das relações de justificação deve poder analisar se isto é uma possibilidade real ou não” (FORST, Rainer. **Justificação e Crítica: perspectivas de uma teoria crítica da política**. Tradução: Denilson Luis Werle. São Paulo: Editora Unesp Digital, 2019, posição 161. E-book Kindle).

¹¹⁵⁷ Ibid., posição 435.

passivos.¹¹⁵⁸ A partir dessa posição, o pensamento humano na justificação das decisões deve ser mantido, em cooperação com os sistemas de Inteligência Artificial judicial, já a Justiça não é uma máquina em si mesma, mas depende do resultado prático para vida a partir das decisões que produz.¹¹⁵⁹

Todavia, é possível ser verificado que ainda há um longo caminho até o desenvolvimento de programas baseados em Inteligência Artificial que consigam realizar justificação de decisões judiciais com grande acurácia. Por exemplo, o projeto *Sinapses* da equipe de tecnologia do Tribunal de Justiça do Estado de Rondônia demonstrou avanços no campo da estruturação de sistema apto a suportar diferentes abordagens de IA no módulo denominado *Cranium* (mediante redes neurais, *machine learning* e lógica bayesiana). Contudo, por experiência prática inerente à presente pesquisa¹¹⁶⁰ os modelos de Inteligência Artificial utilizados no caso ainda não eram capazes de realizar a propositura de conceitos normativos com grau de acurácia sequer mediano, muito pelo fato de que ainda se encontravam em processo de desenvolvimento pelas equipes responsáveis (no TJRO e no CNJ).¹¹⁶¹

Assim, a prática tem demonstrado que, antes de existir tecnologia adequada, é preciso que o método jurídico seja aplicado como modelo para a estruturação do processo de desenvolvimento dos sistemas de IA¹¹⁶², a fim de que soluções de Inteligência Artificial baseadas em *machine learning* sejam um ponto de partida melhor e mais estruturado pela racionalidade jurídica. Isso passa, necessariamente, pela premissa de que o processamento de linguagem natural deve ser projetado como

¹¹⁵⁸ “No entanto, mais do que isso, a questão política sobre quem decide (e como) sobre as estruturas de produção e de distribuição, ou seja, em segundo lugar, com isso ignora-se a questão do poder, como se pudesse existir uma grande máquina distributiva que só precisaríamos programar corretamente. Mas uma tal máquina não apenas não pode existir, visto que assim a justiça não mais poderia ser entendida como uma realização dos próprios sujeitos, como também os transformaria em beneficiários passivos” (Id.).

¹¹⁵⁹ “Por conseguinte, a justiça é uma máquina que produz algo, e unicamente importa o que é produzido e não a vida interna da máquina. Com isso, porém, se perdem os elementos essenciais da justiça social e política, segundo a qual se trata da configuração de um contexto de cooperação política e social, e a primeira questão aqui é como os indivíduos estão envolvidos na produção dos bens materiais e imateriais, de modo que um resultado só pode ser justo quando é produzido sob condições capazes de serem consideradas justas por todos. Seria uma coisa muito boa se existisse um grande Leviatã que distribuisse o maná, mas isso teria pouco a ver com a justiça política e social” (Ibid., posição 564-467).

¹¹⁶⁰ O autor participou de curso de formação em inteligência artificial aplicada ao Judiciário, promovido em setembro de 2019 pelo Conselho Nacional de Justiça e ministrado pela equipe de tecnologia da informação do Tribunal de Justiça do Estado de Rondônia (TJRO).

¹¹⁶¹ Há no âmbito nacional a rede de colaboradores do modelo cooperativo de desenvolvimento de inteligência artificial a ser integrado ao *Sinapses*, modelo este que foi adotado pelo Conselho Nacional de Justiça como política pública do Poder Judiciário.

¹¹⁶² HARTMANN PEIXOTO, Fabiano; SILVA, Roberta Zumblick Martins da. Op. cit., p. 97.

processamento de linguagem jurídica, sem a pretensão de retorno à Jurisprudência dos Conceitos¹¹⁶³, mas sim em busca de um *dataset* que leve em conta os conceitos jurídico-normativos (na proposta de Alexy), muito dependentes da construção dogmática por meio da doutrina¹¹⁶⁴ e dos precedentes judiciais.¹¹⁶⁵ Sendo assim, podem ser aproveitados argumentos dogmáticos (doutrinários), típicos de fase de justificação externa, para a validação de formas de uso de padrão de proposições normativas previstas em lei.¹¹⁶⁶

Nessa linha, no processo de justificação da decisão judicial a fase da justificação interna não é hermeticamente fechada em relação a influências sutilmente advindas da justificação externa, conforme se observa de uma leitura mais detalhada da teoria de Alexy. Em suma, existem pontos de ligação, sendo que a textura aberta da linguagem (uma característica que já havia sido demonstrada por Hart¹¹⁶⁷) ou a decomposição de premissas normativas em proposições linguísticas são exemplos disso. Assim, a utilização de conceitos doutrinários ou trazidos em precedentes judiciais¹¹⁶⁸, desde que devidamente valorados (e aqui os intérpretes humanos são

¹¹⁶³ “A nova Ciência do Direito tem a tendência de ir o mais longe possível na decomposição dos conceitos. E nisso consiste seu mérito. Pois, de fato, da exaustiva compreensão do conteúdo dos conceitos contidos nas proposições jurídicas depende não só a total compreensão do Direito, mas também a segurança na sua aplicação. A decisão final é resultado de um cálculo, cujos são os conceitos jurídicos; o cálculo, evidentemente produzirá um resultado tão mais seguro quanto mais determinado o valor dos fatores. Ao mesmo tempo, é óbvio que só a partir da completa compreensão dos conceitos jurídicos se pode revelar o verdadeiro sistema do Direito, a união interna entre suas proposições” (ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 221-222).

¹¹⁶⁴ Ibid., p. 223-224.

¹¹⁶⁵ Ibid., p. 237-241.

¹¹⁶⁶ “No entanto, da constatação da insuficiência da exclusividade da análise lógica do direito vigente não se infere que a utilização de argumentos sistemático-conceituais na argumentação jurídica seja supérflua ou inclusive nociva. Mais ponderado é considerar que o uso de argumentos sistemático-conceituais, junto com outros argumentos, em especial argumentos práticos do tipo geral, é necessário e racional. Uma das tarefas mais importantes, mas também mais difíceis, de uma teoria do discurso jurídico é a de assinalar tanto os limites como a legitimidade da argumentação sistemático-conceitual” (Ibid., p. 222).

¹¹⁶⁷ “Não apenas no terreno das normas, mas em todos os campos da existência, há um limite, inerente à natureza da linguagem, para a orientação que a linguagem geral pode oferecer. É certo que existem casos claros, que reaparecem constantemente em contextos semelhantes, aos quais as fórmulas gerais são nitidamente aplicáveis (‘Se algo é um veículo, um automóvel o é’), mas haverá também casos aos quais não está claro se elas se aplicam ou não (‘A palavra aqui usada, ‘veículo’, incluirá bicicletas, aviões, patins?’)” (HART, Herbert L. A. Op. cit., p. 164).

¹¹⁶⁸ “Propõe-se, assim, a utilização da técnica do precedente se esse termo for entendido como ‘auto-precedente’, ou seja, como a doutrina ou os critérios adotados pelo próprio juiz ou tribunal para a resolução de casos anteriores. Consequentemente, o respeito ao precedente supõe, nessa concepção, uma exigência de autocongruência, procedimento esse que, em princípio, dirige-se ao passado como fidelidade a casos anteriores, mas que se remete também — e principalmente — ao futuro, neste caso como consciência de que um bom critério de resolução será o que deverá ser aplicado também para resolver casos posteriores, ou seja, o critério que se possa pretender que se transforme em precedente” (CELLA, José Renato Gaziero; WOJCIECHOWSKI, Paola Bianchi. Op. cit.).

necessários), tem o potencial de retirar a questão da justificação da decisão judicial do impasse típico de “problema IA-completo”¹¹⁶⁹ (de muito difícil solução com os limites da tecnologia atual) para uma estruturação que evite o problema da explosão combinatorial. Dessa forma, a utilização de argumentos dogmáticos (doutrinários) e de precedentes tem o potencial para diminuir o risco¹¹⁷⁰ de *human bias*¹¹⁷¹ (enviesamento da IA decorrente de base de dados não representativa da realidade).¹¹⁷²

Sistemas baseados em Inteligência Artificial vêm sendo cada vez mais aplicados ao Direito. Logo, a decisão judicial não está escapando ilesa às mudanças. Contudo, decidir no âmbito do Direito não é como jogar xadrez ou *go* (para citar as vitórias de *Deep Blue* contra Garry Kasparov em 1997 e de *Alpha Go* em face do antes imbatível Lee Sedol em 2016).

Todavia, abordagens como no sistema canadense *ROSS Intelligence*, baseado no processamento de linguagem natural do *Watson*, podem vir a apresentar resultados promissores no futuro, pois já auxiliam profissionais do direito na interpretação de documentos e na propositura de estratégias na solução de problemas. Porém, isso passa muito longe de algo que indique um modelo racional de decisões.

¹¹⁶⁹ “Senso comum e compreensão de linguagem natural também têm se mostrado difíceis. Considera-se com frequência que atingir um desempenho de nível plenamente humano nessas tarefas é um problema “IA-completo”, significando que a dificuldade de solucionar esses problemas se equivale à dificuldade de se construir máquinas com inteligência geral de nível humano. Em outras palavras, se alguém *tivesse* sucesso na criação de uma IA que entendesse a linguagem natural tanto quanto um humano adulto, esse alguém muito provavelmente já teria logrado sucesso na criação de uma IA capaz de fazer tudo mais que a inteligência humana pode fazer, ou estaria muito próximo disso” (BOSTROM, Nick. Op. cit., posição 726-732).

¹¹⁷⁰ “(...) os sistemas jurídicos inteligentes são passíveis de críticas devido ao fato de atenderem exclusivamente ao modo e critérios aplicados pelo seu criador, o programador. Como não é possível, por enquanto, se falar em uma inteligência artificial verdadeiramente autônoma, os softwares existentes operam de maneira condicionada, sempre respondendo aos inputs e outputs pré-determinados e estabelecidos por quem desenvolve o programa. Disso decorre que o resultado das eventuais decisões tomadas pelo computador continuará fortemente influenciado pelos valores, crenças e convicções da pessoa que criou a inteligência artificial, por mais que se busque uma pretensa imparcialidade e superação do subjetivismo” (COSTA, Ramon Silva; OLIVEIRA, Samuel Rodrigues. Pode a Máquina Julgar? Considerações Sobre o Uso de Inteligência Artificial no Processo de Decisão Judicial. **Revista de Argumentação e Hermenêutica Jurídica**, v. 4, n. 2, p. 21-39, jul/dez. 2018).

¹¹⁷¹ SUNSTEIN, Cass R. Algorithms, Correcting Biases.

¹¹⁷² “Há quem espere que a automatização de decisões por computadores represente a superação de questões subjetivas: que tal inteligência por ser artificial seja também imparcial. De fato: a matemática e as ciências exatas que são bases para o desenvolvimento dessas tecnologias são exatas. Mas tanto as soluções de IA são projetadas por pessoas quanto – ainda que se fale em aprendizado de máquina – esse aprendizado se realiza com base na experiência dos algoritmos em um dataset que é produto das relações humanas” (HARTMANN PEIXOTO, Fabiano; SILVA, Roberta Zumblick Martins da. Op. cit., p 102).

Assim, a acurácia dos modelos de decisão baseados em Inteligência Artificial deve ser exigida mediante um modelo de justificação racional inerente à teoria da argumentação jurídica, a fim de se garantir a necessária prestação de contas acerca do método de decisão (art. 93, IX, da Constituição Federal, art. 489, *caput*, II, e §1º, do Código de Processo Civil, e art. 20, §1º, da Lei 13.709/2018), aspecto que a democracia não pode abrir mão.

Desse modo, não é porque existem juízes que descumprem que o dever de fundamentação analítica da decisão (art. 93, IX, da CF e art. 489, §1º, do CPC) que se pode criar um sistema com Inteligência Artificial especializada em decidir de maneira não justificada racionalmente. Nesse caso, a lógica do Direito demonstra que o descumprimento da norma é prova de a que norma existe. Justamente para evitar abusos por parte do julgador, existe a exigência de justificação racional das decisões, o que é inerente à atividade jurisdicional e à própria democracia que se baseia na possibilidade de *accountability* dos atos dos agentes que exercem poder do Estado.

Portanto, é necessária a preocupação com a acurácia da justificação das decisões judiciais auxiliadas pela Inteligência Artificial, a fim de que a argumentação jurídica seja observada como modelo apto a inspirar o desenvolvimento de sistemas de Inteligência Artificial de auxílio à decisão judicial.

4.3 Possibilidades da Justificação Externa na Inteligência Artificial Judicial

4.3.1 Justificação Externa e IA Judicial: Da Pretensão de Correção Normativa à Pretensão de Correção Moral

Esta obra utiliza a teoria da argumentação jurídica como marco teórico, o que acabaria a trazer o leitor para o debate sobre a ponderação de direitos fundamentais através da proporcionalidade.

Isso leva ao delineamento de uma Teoria do Direito para a Inteligência Artificial Judicial num outro nível do debate, o da correção moral típica da fase da justificação externa.

A pretensão de correção, categoria teórica que inspira esta pesquisa, se divide em **pretensão de correção normativa** (no contexto do discurso e razão prática

inerente à argumentação jurídica)¹¹⁷³ e **pretensão de correção moral**¹¹⁷⁴ (que passa pela aplicação de princípios jurídicos e encontra seu ápice no argumento da injustiça¹¹⁷⁵). Nesse sentido, a pretensão de correção engloba os mandamentos de otimização¹¹⁷⁶ advindos do princípio da segurança jurídica e do princípio da justiça, ainda que sejam contravalentes.¹¹⁷⁷ Conforme já trabalhado no subcapítulo 1.6, buscou-se expor o duplo canal de correção nascido da interação comunicacional

¹¹⁷³ “Como ideia reguladora, o conceito de correção não pressupõe que exista para cada questão prática uma resposta correta que deve ser descoberta. Uma única resposta correta é a finalidade a que se deve aspirar. Os participantes em um discurso prático, independentemente de haver uma única resposta correta, devem formular a pretensão de que sua resposta é a única correta. Caso contrário, seriam sem sentido suas afirmações e fundamentações. Isso somente pressupõe haver questões práticas às quais se pode atribuir no discurso uma resposta como a única correta e que não há segurança sobre quais sejam essas questões, de maneira que vale a pena tentar encontrar uma única resposta correta em toda questão. Por isso, a teoria do discurso tem como base uma concepção *absolutamente procedimental da correção*” (ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 271).

¹¹⁷⁴ “O problema nuclear da teoria do discurso é o estabelecimento e justificação do sistema de regras. Ele deve, por um lado, ser tão forte quanto possível para excluir muito tanto quanto possível e, assim, elevar a sua significância de decisão. Por outro, ele deve ser tão fraco quanto possível a fim de ele poder encontrar aprovação tão ampla quanto possível. Sistemas que não mostram nenhum conteúdo moral satisfazem à segunda exigência, não, porém, a primeira. Quem, por exemplo, somente exige a observância das regras da lógica, a verdade das premissas empíricas empregadas e a consideração das regras da racionalidade da finalidade irá encontrar para cada uma dessas exigências facilmente ampla aprovação, seu sistema, porém, sob o pressuposto e participantes com ideias normativas diferentes, mal exclui algo. Além disso, ele não é capaz de esgotar os múltiplos aspectos do conceito de racionalidade.

Os valores, em sentido contrário, da fortidão e fraqueza podem ser trazidos em uma relação ótima quando, ao lado de uma série de outras exigências, sobretudo, são acolhidas as cunhagens distintas da ideia de generalizabilidade como regra do procedimento em $R^p_1, \dots, R^p_n$. O sistema, que nasce desse modo, pode ser designado como ‘kantiano’. Seja aqui partido disto, eu essas regras, como também as regras restantes do discurso, são fundamentáveis sob o pressuposto de um conceito de fundamentação amplo. Razão prática pode ser definida como capacidade de, segundo esse sistema de regras, chegar a conhecimentos práticos” (ALEXY, Robert. **Direito, Razão, Discurso**, p. 83).

¹¹⁷⁵ “Até agora, mostrou-se unicamente que o argumento dos princípios leva a uma conexão necessária entre direito e uma moral qualquer. A objeção leva a crer que isso seria muito pouco. Quando se fala de uma conexão necessária entre direito e moral, geralmente tem-se em mente uma conexão necessária entre o direito e a ou uma moral correta. Isso se aplica especialmente à perspectiva do participante. Na realidade, essa objeção atingiria o não positivista se o argumento dos princípios não conseguisse estabelecer uma conexão necessária qualquer entre direito e a moral correta. Fato de ele conseguir fazê-lo constitui o conteúdo da *tese da correção*. Esta última é o resultado de uma aplicação do argumento da correção nos limites do argumento dos princípios” (ALEXY, Robert. **Conceito e Validade do Direito**, p. 92-93).

¹¹⁷⁶ ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 90.

¹¹⁷⁷ “Poder-se-ia supor que a necessidade da positividade implica o positivismo. Isso, contudo, seria incompatível com a dupla natureza do direito. De acordo com a tese da dupla natureza, o direito possui, ao mesmo tempo, tanto uma dimensão real quanto uma dimensão ideal. A necessidade desse dualismo decorre de dois princípios contravalentes: o princípio da segurança jurídica (Rechtssicherheit) e o princípio da justiça. Em virtude da fraqueza da dimensão ideal, a pretensão de correção compreende ambos os princípios. Isso implica que ambos estão necessariamente conectados ao direito. O princípio da segurança jurídica exige um compromisso com aquilo que é autoritativamente produzido e socialmente eficaz. O princípio da justiça exige que a decisão seja moralmente correta. Nenhum dos dois pode suprimir completamente o outro” (ALEXY, Robert. **Princípios Formais**. Rio de Janeiro: Forense, 2018, p. 87-88. E-book Kindle).

humano-máquina potencializado e aprimorado pelos aspectos inerentes à correção normativa.

Por outro lado, a correção moral tem elementos mais desafiadores ao intérprete do Direito, especialmente quando se busca a aplicação dos princípios jurídicos de natureza constitucional, oriundos de normas de direitos humanos ou advindos de normas infraconstitucionais.

Nesse caminho, de correlação entre correção normativa e correção moral, a aplicação de normas com estrutura de princípios é necessária, haja vista que o Direito brasileiro conjuga regras¹¹⁷⁸ e princípios¹¹⁷⁹. Assim, a estrutura linguística de princípios, que “representam razões que podem ser afastadas por razões antagônicas”¹¹⁸⁰, passa a ser enfocada no presente texto, levando-se em consideração o recorte que o estudo da Inteligência Artificial Judicial traz ao tema e sua relação com a obra aqui desenvolvida.

A partir destas premissas, percebe-se que a proporcionalidade remete à necessidade de aplicação de dois ou mais princípios em interação num caso concreto¹¹⁸¹, o que se dá através de um sopesamento entre princípios que, embora *prima facie*¹¹⁸² sejam abstratos, resultam em uma representação de solução jurídica para um caso concreto.¹¹⁸³

Assim, como visto anteriormente no subcapítulo 4.1, as representações podem ser programadas em sistemas computacionais de IA jurídica mediante fatores (*factors*) ou conceitos jurídicos intermediados (*ILC's*). Por exemplo: imagine o já comentado caso do personagem José (vide o subcapítulo 1.6), nele poderia existir uma

¹¹⁷⁸ “Como as regras exigem que seja feito exatamente aquilo que elas ordenam, elas têm uma determinação da extensão do seu conteúdo no âmbito das possibilidades jurídicas e fáticas. Essa determinação pode falhar diante de impossibilidades jurídicas e fáticas; mas, se isso não ocorrer, então, vale definitivamente aquilo que a regra prescreve” (ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 104).

¹¹⁷⁹ “Princípios exigem que algo seja realizado na maior medida possível dentro das possibilidades jurídicas e fáticas existentes. Nesse sentido, eles não contêm um *mandamento definitivo*, mas apenas *prima facie*” (Ibid., p. 103-104).

¹¹⁸⁰ Ibid., p. 104.

¹¹⁸¹ E aqui se remete, mais uma vez, aos conceitos de suporte fático abstrato e suporte fático concreto, bem como à sua relação com os fatores (*factors*) e com os conceitos jurídicos intermediados (*ILC's*) necessários à representação do Direito nos sistemas de inteligência artificial.

¹¹⁸² “Princípios são, ao contrário, sempre razões *prima facie*. Isoladamente considerados, eles estabelecem apenas direitos *prima facie*” (ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 107).

¹¹⁸³ “O ‘conflito’ deve, ao contrário, ser resolvido ‘por meio de um sopesamento entre interesses conflitantes’. O objetivo desse sopesamento é definir qual dos interesses – que *abstratamente estão no mesmo nível* – tem *maior peso no caso concreto*” (Ibid., p. 95).

representação da proporcionalidade no caso de conflito entre as pretensões do credor (baseado nos princípios de segurança jurídica e de proteção do cumprimento do contrato) e do devedor (baseado na dignidade e na justiça), de maneira que o sistema passasse a refletir a derroabilidade da Súmula 486 do Superior Tribunal de Justiça frente aos argumentos previamente ponderados.

Dessa forma, há pretensão de correção normativa e pretensão de correção moral em uma representação advinda da aplicação da proporcionalidade. Ao final, a representação da proporcionalidade é uma regra que passa a integrar o sistema de IA judicial, que sempre pode ser desafiada pelo julgador na busca de restabelecer a correção para o caso concreto, pois os princípios, como mandamentos de otimização *prima facie*, são “aplicados na medida das possibilidades fáticas e jurídicas de sua realização”.¹¹⁸⁴

Além disso, destaca-se que no nível abstrato, não existe uma prévia prevalência de um princípio em relação ao outro¹¹⁸⁵, o que inviabiliza o estabelecimento de cláusulas de exceção que sejam totalmente vinculativas ao intérprete. Tal realidade demanda o ininterrupto funcionamento do duplo canal de correção na decisão judicial auxiliada por IA, o que torna o elemento humano uma figura perene nesse processo.

Logo, é possibilitada uma “relação de precedência condicionada entre os princípios, com base nas circunstâncias do caso concreto”¹¹⁸⁶, o que possibilita o estabelecimento de certas circunstâncias do caso concreto que podem ser representadas por fatores (*factors*) ou podem ser pressupostas em *ILC's* (Conceitos Jurídicos Intermediados).

Dessa maneira, a programação de sistemas de representação de conhecimento e, em certa medida, a escolha¹¹⁸⁷ de modelos de IA baseados em

¹¹⁸⁴ Id.

¹¹⁸⁵ “Portanto, isoladamente considerados, ambos os princípios conduzem a uma contradição. Isso significa, por sua vez, que um princípio restringe as possibilidades jurídicas e fáticas de realização do outro. Essa situação não é resolvida com a declaração de invalidade de um dos princípios e com a sua consequente eliminação do ordenamento jurídico. Ela tampouco é resolvida por meio da introdução de uma exceção a um dos princípios, que seria considerado em todos os casos futuros, como uma regra que ou é realizada, ou não é” (Ibid., p. 96).

¹¹⁸⁶ “A solução para essa colisão consiste no estabelecimento de uma relação de precedência condicionada entre os princípios, com base nas circunstâncias do caso concreto. Levando-se em consideração o caso concreto, o estabelecimento de relações de precedências condicionadas consiste na fixação de condições sob as quais um princípio tem precedência em face do outro. Sob outras condições, é possível que a questão da precedência seja resolvida de forma contrária” (Id.).

¹¹⁸⁷ O desenvolvimento de sistemas de baseados em aprendizado de máquina envolve escolhas da equipe de desenvolvimento.

machine learning deve passar por uma análise prévia de proporcionalidade que é necessária para o prévio estabelecimento das condições de precedência entre princípios¹¹⁸⁸, pois a programação de sistemas de IA judicial trata de antever soluções para futuros casos concretos (ainda que a derrotabilidade do modelo computacional possa ser acionada pelo juiz ciborgue). A IA judicial busca “acertar o alvo”, num esforço do duplo canal de correção humano-máquina, para auxiliar no procedimento argumentativo de julgamento, mas o julgador humano pode justificadamente deixar de aplicar a ponderação representada no sistema.

Portanto, descabe o uso de sistemas de IA para o estabelecimento de relações incondicionadas de precedência de um princípio constitucional em relação a outro, uma vez que esse tipo de precedência é juridicamente ilegítimo ao não refletir as circunstâncias concretas de aplicação.¹¹⁸⁹ Por outro lado, podem ser estabelecidas relações de precedência abstrata entre princípios de diferentes hierarquias normativas (um princípio constitucional tem precedência sobre princípios legais). Ademais, relações de precedência concreta¹¹⁹⁰, para suportes fáticos concretos, podem ser simuladas em relação a suportes fáticos hipotéticos presentes no sistema de auxílio à decisão judicial.

Em suma, a representação de suportes fáticos concretos que já ocorreram no passado (precedentes) ou a representação de suportes fáticos que ainda não ocorreram, mas que são uma possibilidade, são formas de se estabelecer suporte fático voltado à programação de sistemas de computação jurídica ou à modelagem de IA judicial baseada em aprendizado de máquina.

¹¹⁸⁸ “Essa colisão pode ser resolvida ou por meio de uma relação de precedência incondicionada ou por meio de do estabelecimento de uma relação de precedência condicionada. Como símbolo para a relação de precedência deve ser usado o sinal **P**. Para as condições sob as quais um princípio tem precedência em face do outro será utilizado o sinal **C**” (Ibid., p. 96-97).

¹¹⁸⁹ “Há, então, quatro possibilidades de decisão do caso a partir da solução de uma colisão entre princípios:

- (1) $P_1 \mathbf{P} P_2$.
- (2) $P_2 \mathbf{P} P_1$.
- (3) $(P_1 \mathbf{P} P_2) C$.
- (4) $(P_2 \mathbf{P} P_1) C$.

(1) e (2) são relações incondicionadas de precedência. É possível falar também em relações de precedências ‘abstratas’ ou ‘absolutas’. O Tribunal Constitucional Federal excluiu a possibilidade dessa forma de relação de precedência com a afirmação: ‘nenhum desses interesses goza, em si mesmo, de precedência sobre o outro. Essa afirmação vale de forma mais geral para as colisões entre princípios de direito constitucional’ (Ibid., p. 97).

¹¹⁹⁰ “Sobra, portanto, apenas a possibilidade de uma relação condicionada, ou, como também se pode dizer, de uma relação de precedência concreta ou relativa, como as indicadas sob os ns. (3) e (4). A questão decisiva é, portanto, sob quais condições qual princípio deve prevalecer e qual deve ceder” (Id.).

Nessa relação de precedência condicionada há uma resposta simples, na qual dado um determinado suporte fático concreto, um princípio P_1 supera um princípio colidente P_2 , desde que a partir de determinado suporte fático concreto existam “razões suficientes para que P_1 prevaleça sobre P_2 sob as condições C , presentes nesse caso concreto”.¹¹⁹¹

Dessa maneira, quando considerado o mesmo nível hierárquico entre princípios (princípios advindos de direitos fundamentais, por exemplo) não podem ser estabelecidas relações incondicionais de precedência.¹¹⁹² Por outro lado, consideradas as legítimas regras¹¹⁹³ constitucionais de distribuição de competência legislativa (art. 22, art. 24, art. 25, §1º e §3º, e art. 30, I e II, da Constituição Federal¹¹⁹⁴) e a supremacia da Constituição, podem ser estabelecidas relações incondicionais de precedência, mediante a definição de um suporte fático hipotético parcial que contenha apenas a matéria de competência¹¹⁹⁵, de maneira que o sistema de IA judicial sugira o(s) princípio(s) aplicável(is) ao caso concreto. Assim, um sistema de IA judicial voltado ao auxílio do julgamento pode identificar, por exemplo, se um caso de direito privado envolve uma relação de precedência incondicionada entre a proteção ao consumidor (art. 5º, XXXII, e art. 170, V, da Constituição Federal) e a autonomia da vontade (art. 421 do Código Civil).

Sob certas condições de um suporte fático hipotético (C_1) presente na IA judicial, o princípio P_1 prevalece sobre o princípio P_2 .¹¹⁹⁶ Já sob outras condições de um outro suporte fático hipotético (C_2), o princípio P_2 prevalece sobre o princípio P_1 .¹¹⁹⁷ Nesse cenário, o estabelecimento de uma relação de precedência condicionada entre princípios para o caso de determinado suporte fático hipotético é a representação de uma regra.¹¹⁹⁸

¹¹⁹¹ Id.

¹¹⁹² Id.

¹¹⁹³ Ibid., p. 106.

¹¹⁹⁴ Estes são os dispositivos que estabelecem as principais regras de distribuição de competência legislativa, o que não exclui a análise de outros dispositivos constitucionais que se relacionam com o tema.

¹¹⁹⁵ “A *differentia specifica* dos princípios materiais é que seus objetos de otimização são determinados conteúdos, como, por exemplo, a vida, a liberdade de expressão, o mínimo existencial e a proteção do meio ambiente. Em contraste, os objetos de otimização de princípios formais são decisões jurídicas, independentemente de seus conteúdos. Princípios formais exigem que a autoridade de normas expedidas devidamente (em conformidade com o ordenamento jurídico) e socialmente eficazes seja otimizada” (ALEXY, Robert. **Princípios Formais**, p. 9).

¹¹⁹⁶ $(P_1 \text{ P } P_2) C_1$.

¹¹⁹⁷ $(P_2 \text{ P } P_1) C_2$.

¹¹⁹⁸ “ C é a condição de uma relação de precedência. Na formulação da regra:

(5) Se uma ação h preenche C , então, h é proibida sob o ponto de vista dos direitos fundamentais,

Desse modo, a regra advinda de princípios aplicados a um suporte fático hipotético funciona com base na programação (um enunciado) de operador condicional “se, então” (“*if, then*”) de representação de relação de precedência condicionada que prevê a consequência da aplicação do princípio prevalente.¹¹⁹⁹

Assim, há a criação de uma regra (K), na qual caso o princípio P_1 tenha precedência em relação do princípio P_2 sob as condições C , ou seja $(P_1 \mathbf{P} P_2) C$, e caso ocorram as condições C (hipótese de precedência do princípio P_1), haverá a consequência jurídica R . Dessa forma, para a Inteligência Artificial judicial voltada à representação de princípios em colisão, a regra computacionalmente programada ou modelada na IA decorre de **C como suporte fático hipotético** e tem R como a consequência jurídica¹²⁰⁰, ou seja: $C \rightarrow R$.¹²⁰¹

Conquanto a “lei de colisão” defendida por Alexy, que enxerga os princípios como mandamentos de otimização, considere que não existe relação absoluta de precedência entre princípios e que a totalidade do suporte fático abstrato (referência a ações e situações) não é previamente quantificável¹²⁰², evidencia-se a utilidade da representação de relações de precedência condicional, $(P_1 \mathbf{P} P_2) C_1$ ou $(P_2 \mathbf{P} P_1) C_2$, para estabelecer de maneira legítima e explicável, que, caso verificado determinado **suporte fático hipotético** (*input*), existirão **determinadas consequências jurídicas** (*output*): $C_n \rightarrow R_n$. Ao final deste procedimento interpretativo, as diversas consequências jurídicas (R_n) para cada um dos suportes fáticos hipotéticos (C_n) podem resultar em regras de proibição, obrigação, permissão ou facultatividade ($C_n \rightarrow R_{PH} \vee R_O \vee R_P \vee R_F$).

A modelagem em lógica proposicional é utilizada aqui apenas em nível abstrato, na medida em que determinado suporte fático hipotético pode levar a uma ou mais regras, dependendo do microssistema de IA judicial em que esteja representado. Por exemplo, se o juiz deixa de oportunizar contraditório à parte para que se manifeste sobre um documento utilizado como fundamento da decisão

C é o pressuposto do suporte fático de uma regra” (ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 98).

¹¹⁹⁹ “Por conseguinte, de um enunciado de preferência acerca de uma relação condicionada de precedência decorre uma regra, que, diante da presença da condição de precedência, prescreve a consequência jurídica do princípio prevalente” (Ibid., p. 99).

¹²⁰⁰ Id.

¹²⁰¹ “(K’) As condições sob as quais um princípio tem precedência em face de outro constituem o suporte fático de uma regra que expressa a consequência jurídica do princípio que tem precedência” (Id.).

¹²⁰² Id.

judicial¹²⁰³ (suporte fático hipotético C que pode ser representado por um *factor*), então decorrem as seguintes consequências jurídicas: uma facultatividade da parte prejudicada em recorrer, haja vista o interesse processual (R_F); uma obrigação ao Tribunal no sentido de anular a sentença (R_O), ou, por outro lado, uma proibição de manutenção da sentença (R_{PH})¹²⁰⁴; e uma permissão ao Tribunal para que antecipe de tutela recursal (R_P). A depender do objetivo e do alcance do sistema de IA judicial, qualquer das consequências pode vir a ser representada através de programação computacional ou modelos baseados em aprendizado de máquina.¹²⁰⁵

Além disso, Alexy destaca que, a partir da condição de precedência de um princípio em relação a outro¹²⁰⁶, “o suporte fático da regra corresponde ao enunciado de preferência segundo a lei de colisão”¹²⁰⁷ de maneira a incluir o estabelecimento de regras de exceção ao princípio restringido no procedimento de proporcionalidade.¹²⁰⁸

Observe-se que, para a decisão que “ocorre na terceira etapa”¹²⁰⁹ do procedimento da proporcionalidade (na fase da proporcionalidade em sentido estrito), a análise computacional do **suporte fático hipotético** presente na IA judicial pode tratar a condição de precedência¹²¹⁰ de maneira concentrada (o que é possível se fazer através de um conceito jurídico intermediado, ou *ILC*) ou pode, por outro lado, levar em consideração cada um dos elementos fáticos (fatores ou *factors*) que compõem a condição de precedência C , o que Alexy resume da seguinte forma: T_1 e T_2 e T_3 e $T_4 \rightarrow R$.¹²¹¹ Por outro lado, todas as condições T_{1-4} podem ser incluídas num único *fator*, caso esta seja a escolha de representação.

¹²⁰³ Uma inobservância ao princípio do contraditório (art. 5º, LV, da Constituição Federal) sob a alegação de prevalência do princípio da razoável duração do processo (art. 5º, LXXVIII, da Constituição Federal).

¹²⁰⁴ É um caso onde a proibição e a obrigação são intercambiáveis, a depender da forma de representação.

¹²⁰⁵ Por exemplo, um modelo de IA que auxilie na análise da probabilidade de admissibilidade recursal se voltará para aspectos distintos de um modelo que auxilie a fundamentação do mérito do acórdão.

¹²⁰⁶ ($P_1 \mathbf{P} P_2$), por exemplo.

¹²⁰⁷ ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 101.

¹²⁰⁸ *Id.*

¹²⁰⁹ *Id.*

¹²¹⁰ “O que aqui interessa é somente que a aceitação de uma carga argumentativa em favor de determinados princípios não iguala seu caráter *prima facie* ao das regras. Mesmo uma regra sobre ônus argumentativo não exclui a necessidade de definir as condições de precedência no caso concreto. Ela tem como consequência apenas a necessidade de se dar precedência a um princípio em relação a outro caso haja razões equivalentes em favor de ambos ou em caso de dúvida” (*Ibid.*, p. 106).

¹²¹¹ “Nesse sentido, vale o enunciado de preferência ($P_2 \mathbf{P} P_1$) C_2 . C_2 é composto por quatro condições (repetição/ausência de interesse atual pela informação/ grave crime/risco à ressocialização). A regra $C_2 \rightarrow R$, que corresponde ao enunciado de preferência, é uma regra com quatro atributos de suporte fático, com a seguinte estrutura:

(6) T_1 e T_2 e T_3 e $T_4 \rightarrow R$.

No mais, fica evidenciado que é possível se pensar no uso da IA para buscar soluções jurídicas para suportes fáticos previamente definidos. Ou seja, representações de suportes fáticos concretos, porém que têm aplicabilidade prática que supera a mera abstração. Por outro lado, considerada a entropia jurídica, ainda não se mostra possível¹²¹² o desenvolvimento de IA jurídica que crie suportes fáticos hipotéticos e aplique o direito.

A partir dessa realidade, levando-se em conta as características inerentes à aplicação de princípios jurídicos, ainda que sistemas de Inteligência Artificial judicial (baseados em *factors*, *ILC's* ou em modelos de *machine learning*) possam chegar à **representação de regras resultantes da aplicação de proporcionalidade para diversos suportes fáticos hipotéticos** (o que demandaria bastante esforço da equipe interdisciplinar de desenvolvimento, especialmente dos assessores-desenvolvedores e dos magistrados), tais regras são soluções que sempre sustentam grau de provisoriedade¹²¹³ e que podem ser superáveis mediante argumentação jurídica ostensiva. Essa superação da regra representada em sistema computacional deve se dar nas seguintes hipóteses: **a)** o suporte fático hipotético (presente no sistema de IA judicial) não corresponde exatamente ao suporte fático concreto; e **b)** apresentação outros elementos jurídicos (um novo princípio ou uma regra que não foram considerados, por exemplo) que influenciam a ponderação pela proporcionalidade ou não recomendam seu uso. A partir disso, o aspecto dinâmico da argumentação jurídica sempre pode pulsar num ritmo embalado pela entropia jurídica, pois mesmo “em um discurso ideal potencialmente infinito nunca é certo se um consenso, uma vez obtido, é terminante ou definitivo”.¹²¹⁴

Talvez uma IA forte com inteligência geral consiga um dia simular um mundo jurídico, desde os suportes fáticos abstratos e concretos até as regras inferidas de

Ou seja: uma notícia repetida (T_1), não revestida de interesse atual pela informação (T_2), sobre um grave crime (T_3), e que põe em risco a ressocialização do autor (T_4), é proibida do ponto de vista dos direitos fundamentais” (Ibid., p. 102).

¹²¹² O avanço das Inteligências Artificiais generativas tem sido rápido no início de 2023 (data de escrita desta nota de rodapé), sendo que é possível se cogitar a aplicação dessa espécie de IA para a simulação de suportes fáticos hipotéticos para o Direito. Contudo, o tema necessita de maior desenvolvimento e maturação junto à Ciência da Computação e, na sequência, no Direito, motivo pelo qual seria precipitado trazê-lo neste momento para o contexto deste subcapítulo. Todavia, o subcapítulo 5.7 faz uma abordagem possível do tema, neste momento histórico (início de 2023), dentro do contexto maior do aprendizado de máquina profundo (*deep learning*).

¹²¹³ E a entropia jurídica, como já comentado no subcapítulo 1.6, possibilita a reanálise de soluções antes tidas como “definitivas”.

¹²¹⁴ ALEXY, Robert. **Direito, Razão, Discurso**, p. 92.

princípios e da relação de proporcionalidade entre eles. Seria uma IA de pouca utilidade prática para humanos, que são racionais na medida de sua irracionalidade (vide o subcapítulo 3.4). Por outro lado, simular isso de maneira controlada (talvez em um metaverso¹²¹⁵) poderia ser um experimento computacional e jurídico bastante interessante para a Ciência, o que fica como um desafio para os presentes e futuros juristas-desenvolvedores.

Nesse caso, a referência ao universo cinematográfico de *Matrix* mais uma vez ajudará o leitor a compreender o gigantismo (e os riscos) dessa empreitada, na qual o personagem *Arquiteto* cautelosamente lembra do “seu fracasso monumental”.¹²¹⁶ Nesse sentido, ao final, ainda que um sistema de IA judicial fosse plenamente adaptável aos novos suportes fáticos concretos, isso ocorreria como resposta ao tridente da entropia jurídica. O sistema é supostamente “perfeito” porque se adapta a tudo (e, de alguma forma, prevê tudo) ou, no outro lado do espelho, pouco prevê e apenas se adapta com base na variabilidade entrópica da realidade? Um Direito perfeito que causasse qualquer mutação em si mesmo (em seus sistemas, princípios e regras) para “prever” a si próprio seria irrealizável: eis o paradoxo do Direito perfeito. Adaptabilidade total e previsibilidade divina são sonhos irrealizáveis. Justiça e segurança jurídica imperfeitas vivem nas relações humanas igualmente imperfeitas, não em um sistema computacional (ou social), por mais abrangente que ele busque ser.

Dessa maneira, observa-se que não há sentido na criação de suporte fático hipotético sem um contexto de aplicação de um sistema normativo, o que denota a representação de regras decorrentes de individualização de princípios (inclusive de relações de precedência no procedimento de proporcionalidade) como a forma mais adequada (e humanamente compreensível) de desenvolvimento da IA judicial.

¹²¹⁵ A parte boa do metaverso ser o novo *hype* na cabeça do aceleracionistas e futuristas é que a presente pesquisa ganhou um pouco de paz epistêmica, deixando de ser necessário responder às cotidianas ideias desconectadas da realidade que surgem (e a obra se preocupou, nos últimos dois anos de escrita, com as discussões sobre IA jurídica conectadas à realidade). Entre dezembro de 2018 (início da escrita do pré-projeto desta obra) e o momento em que são escritas estas linhas (21/11/2022) muita coisa aconteceu em matéria de Inteligência Artificial e, especialmente, relacionada à sua aplicação ao Direito. Esta nota de rodapé é somente “uma página” do diário destes tempos, a fim de incluir o leitor na entropia relacionada à escrita deste estudo.

¹²¹⁶ THE MATRIX Reloaded.

Desse modo, a regra que nasce da relação $C \rightarrow R$ inerente ao procedimento da proporcionalidade¹²¹⁷ (considerados os testes adequação, necessidade e proporcionalidade em sentido estrito) se constitui em si mesma numa norma de direito fundamental atribuída.¹²¹⁸ Em síntese, dos princípios aplicados ao suporte fático representado no sistema computacional, surgem regras para o auxílio à tomada de decisão judicial através de Inteligência Artificial.

Nesse sentido, mesmo a aplicação de regras pelo sistema de IA pode ser influenciada pelo estabelecimento de cláusulas de exceção advindas da individualização de princípios para um determinado suporte fático hipotético¹²¹⁹, ainda que novas exceções possam surgir a partir do elemento humano inerente ao duplo canal de correção.

Por outro lado, nesse ponto, deve ser desmistificada uma lenda surgida do senso comum de parcela dos operadores do Direito, a qual pressupõe equivocadamente que um princípio sempre supera uma regra de igual estatura hierárquica.¹²²⁰ Esta ideia é incorreta, pois conforme destaca Alexy:

“Já uma regra não é superada pura e simplesmente quando se atribui, no caso concreto, um peso maior ao princípio contrário ao princípio que sustenta a regra. É necessário que sejam superados também aqueles princípios que estabelecem que as regras que tenham sido criadas pelas autoridades legitimadas para tanto devem ser seguidas e que não se deve relativizar sem motivos uma prática estabelecida” (destaque não original).¹²²¹

¹²¹⁷ “Decisões sobre direitos pressupõe a identificação de direitos definitivos. O caminho que vai do princípio, isto é, do direito *prima facie*, até o direito definitivo passa pela definição da relação de preferência. Mas a definição de uma preferência é, segundo a lei da colisão, a definição de uma regra”. (ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 107).

¹²¹⁸ “Uma norma de direito fundamental atribuída é uma norma para cuja atribuição é possível uma correta fundamentação referida a direitos fundamentais. Se é possível uma correta fundamentação referida a direitos fundamentais para a norma que se acaba de apresentar – algo que aqui se pressupõe –, então, ela é uma norma de direito fundamental. Ao mesmo tempo ela é uma regra, à qual se podem subsumir os elementos do caso concreto, como se fosse uma norma positiva. Foi o que aconteceu no caso Lebach. Diante disso, pode-se afirmar: como resultado de todo sopesamento que seja correto do ponto de vista dos direitos fundamentais pode ser formulada uma norma de direito fundamental atribuída, que tem estrutura de uma regra e à qual o caso pode ser subsumido” (Ibid., p. 102).

¹²¹⁹ “A introdução de uma cláusula de exceção pode ocorrer em virtude de um princípio. Ao contrário do que sustenta Dworkin, as cláusulas de exceção introduzidas em virtude de princípios não são nem mesmo teoricamente enumeráveis. Nunca é possível ter certeza de que, em um novo caso, não será necessária a introdução de uma nova cláusula de exceção” (Ibid., p. 104).

¹²²⁰ Um princípio do Código de Processo Civil superaria uma regra processual, por exemplo.

¹²²¹ ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 105.

Em sentido complementar, Alexy evidencia a determinabilidade *prima facie* das regras, pois essas se baseiam em “decisões tomadas pelas autoridades legitimadas para tanto ou decorrentes de uma prática reiterada” [o que] “continua a ser algo fundamentalmente diferente e muito mais forte”¹²²², quando comparadas ao caráter *prima facie* dos princípios.

Ou seja, no contexto brasileiro a derrotabilidade de regras a partir de princípios constitucionais encontra limites argumentativos e institucionais, pois as regras têm como alicerce princípios de igual ou superior estatura¹²²³, sendo inadequada a sua superação, pura e simples, mediante mera alegação de um princípio contraposto. Isso demonstra a força do Direito através de suas regras constitucionais escritas¹²²⁴, o que é parte do *corpus* primário que deve ser considerado para o desenvolvimento de sistemas de Inteligência Artificial judicial, mesmo diante da possibilidade de estabelecimento de regras por atribuição de princípio aplicados a determinado suporte fático hipotético. Nesse sentido, a programação padrão, para casos de análise *prima facie* em suporte fático abstrato, deve ser que **uma regra supera um princípio**, pois os princípios “são sempre razões *prima facie* e regras são, se não houver o estabelecimento de alguma exceção, razões definitivas”.¹²²⁵

Outro aspecto que deve ser considerado para o desenvolvimento de sistema IA judicial se dá no sentido de que é possível prever “uma carga argumentativa a favor de interesses individuais e contra interesses coletivos”.¹²²⁶ Dessa forma, o ônus de superação argumentativa de direitos individuais¹²²⁷ em prol de princípios que se

¹²²² Ibid., p. 106.

¹²²³ “Tais princípios devem ser denominados ‘princípios formais’. Em um ordenamento jurídico, quanto mais peso se atribui aos princípios formais, tanto mais forte será o caráter *prima facie* de suas regras. Somente quando se deixa de atribuir algum peso a esse tipo de princípios – o que teria como consequência o fim da validade das regras enquanto regras – é que regras e princípios passam a ter o mesmo caráter *prima facie*” (Ibid., p. 105).

¹²²⁴ “O outro lado é que, mesmo diante de um fortalecimento de seu caráter *prima facie*, os princípios não obtêm um caráter *prima facie* como o das regras. O caráter *prima facie* dos princípios pode ser fortalecido por meio da introdução de uma carga argumentativa a favor de determinados princípios ou de determinadas classes de princípios” (Ibid., p. 105-106).

¹²²⁵ Ibid., p. 106.

¹²²⁶ Id.

¹²²⁷ “A primazia-*prima facie* não exclui o conter de direitos individuais por bens coletivos. Ela pede somente que, nisso, falem fundamentos mais fortes em favor da solução exigida por bens coletivos que para a exigida por direitos individuais. Com isso, o sistema normativo ganha, sem dúvida, uma tendência individualista, ela, porém, é tão fraca que ela não iria bastar a um representante de uma teoria política rigorosamente individualista, o que, adicionalmente, mostra quão pouco sustentável é a objeção do individualismo. A direção da primazia-*prima facie* é ignorada quando ela não é interpretada como primazia de todos os direitos pertencentes à classe dos direitos individuais perante bens coletivos, mas como primazia somente de uma classe parcial desses direitos, por exemplo, somente como primazia dos direitos de liberdade” (ALEXY, Robert. **Direito, Razão, Discurso**, p. 198).

baseiam em interesse coletivo deve sempre ser atribuído ao julgador humano, cabendo ao sistema de IA judicial atuar como uma espécie de *nudge* (vide subcapítulo 3.4) ao lembrar o julgador da primazia *prima facie* dos direitos individuais previstos na Constituição Federal, em seus diversos aspectos, cabendo apenas ao julgador humano o ônus argumentativo da superação de um direito fundamental individual em colisão com outras normas com estrutura de princípio constitucional. A partir disso, a programação ou a modelagem de Inteligência Artificial jurídica devem se pautar no sentido que **direitos individuais fundamentais tem primazia *prima facie* em relação aos direitos coletivos quando considerados os suportes fáticos abstratos**, salvo a superação argumentativa humana através do procedimento da proporcionalidade em determinado suporte fático concreto. Em suma, os direitos fundamentais também atuam como regras sobre ônus argumentativo.¹²²⁸

Dessa forma, no contexto da lógica de estabelecimento de proposições normativas¹²²⁹ sobre o que é proibido, obrigatório, permitido ou facultativo¹²³⁰, o desenvolvimento de IA judicial para tratar de representações feitas através de suporte fático hipotético se dá no contexto de que regras “podem ser também razões para outras regras e princípios podem também ser razões para decisões concretas”¹²³¹, uma vez que regras e princípios são razões para ações.¹²³² Ou seja, os princípios podem ser base sobre juízos a respeito de casos hipotéticos no desenvolvimento de Inteligência Artificial de apoio à decisão judicial. Nesse caso, os princípios atuam como razões para as regras¹²³³ construídas no sistema de IA judicial.

Desse modo, percebe-se que no desenvolvimento de sistema de IA judicial, uma regra, nascida da aplicação de um princípio¹²³⁴ a um suporte fático representado na Inteligência Artificial, estabelece um juízo de dever-ser que passará de hipótese a

¹²²⁸ ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 106.

¹²²⁹ BULYGIN, Eugenio. **Lógica deóntica, normas y proposiciones normativas**, p. 52

¹²³⁰ “Na Ciência do Direito são formulados juízos sobre o que é devido, o que é proibido e o que permitido, e o juiz decide exatamente sobre isso. O conceito semântico de norma é desenvolvido para essa tarefa. Se se compreendem regras e princípios como razões para normas, a relação de fundamentação fica limitada a objetos de uma categoria, o que facilita seu manejo e, sobretudo, sua análise lógica” (ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 107).

¹²³¹ Id.

¹²³² “Aqui, ao contrário, regras e princípios devem ser considerados como razões para normas. O fosso entre ambas as concepções é, no entanto, menos profundo do que pode parecer pois, se regras e princípios são razões para normas, eles são também indiretamente razões para ações” (Id.).

¹²³³ “Ainda assim a caracterização dos princípios como razões para regras indica um ponto acertado” (Id.).

¹²³⁴ “Nesse sentido, é possível afirmar que sempre que um princípio for, em última análise, uma razão decisiva para um juízo concreto de dever-ser, então, esse princípio é o fundamento de uma regra, que representa uma razão definitiva para esse juízo concreto” (Id.).

juízo definitivo no exato momento em que o juiz ciborgue (através de seu duplo canal de correção) ultrapassa a análise do suporte fático hipotético e passa ao suporte fático concreto.¹²³⁵

A partir dessas considerações, ainda que seja possível se pensar em sistemas de IA voltados à análise de proporcionalidade (adequação, necessidade e proporcionalidade em sentido estrito) no caso de **conflito de princípios em um suporte fático concreto**, para a atual fase de desenvolvimento de modelos de IA judicial não é eticamente aconselhável a tentativa de implementação de sistemas voltados à realização completa do processo argumentativo da proporcionalidade, o qual deve ser reservado ao julgador humano. O referido arranjo epistêmico supera ressalvas ao procedimento argumentativo da proporcionalidade através de IA¹²³⁶ ao manter controle humano e possibilitar explicabilidade no funcionamento do duplo canal de correção.

Talvez a utilização do aprendizado de máquina profundo (*deep learning*), a exemplo do que foi feito no desenvolvimento da *Democratic AI*¹²³⁷, possa, um dia, levar à criação de sistemas que tenham performance para a realização de análise de proporcionalidade para inúmeros suportes fáticos concretos não previamente modelados (*machine learning* no paradigma da Ciência de Dados) ou programados (representação do conhecimento). Contudo, isso se aproxima mais de uma tentativa de visão do futuro do que de um relato da realidade atual da relação entre Teoria do Direito, Direito Constitucional e Inteligência Artificial.

Por esses motivos, de maneira mais atrelada à realidade e ao presente em que se vive (ainda que ele seja sempre fugaz), a representação ou a modelagem de condições em termos de **suportes fáticos representados no sistema de IA** se mostram como tecnologicamente possíveis e juridicamente legítimos para a inclusão de Inteligência Artificial na aplicação de princípios jurídicos e, por consequência, na técnica da proporcionalidade no caso de colisão entre princípios (nos limites aqui

¹²³⁵ “Se uma regra é uma razão para um determinado juízo concreto – o que ocorre quando ela é válida, aplicável e infensa a exceções –, então, ela é uma razão definitiva” (Ibid., p. 108).

¹²³⁶ ARAÚJO, Erik da Silva e; SIMIONI, Rafael Larazzotto. A Inteligência Artificial e Ponderação: o Procedimentalismo de Robert Alexy na Forma de Algoritmo. **Revista Jurídica Eletrônica da Universidade Federal do Piauí**, v. 8, n. 1, jan/jun 2021.

¹²³⁷ KOSTER, Raphael; BALAGUER, Jan; TACCHETTI, Andrea *et al.* Human-centred mechanism design with Democratic AI. **Nature Human Behaviour**, v. 6, p. 1398-1407, 2022. DOI: 10.1038/s41562-022-01383-x.

defendidos), pois se baseia no controle permanente do julgador e de sua equipe de desenvolvimento.

Portanto, a proteção de direitos fundamentais em sistemas de IA judicial passa, quando cada direito é individualmente considerado, pelo dever de ser preenchido o suporte fático em sentido estrito (o que ocorre pelo suporte fático hipotético **representado no sistema de IA judicial**) e não pelo não preenchimento de sua cláusula de restrição.¹²³⁸ Desse modo, ao se levar em conta o suporte fático e a restrição, podem ser representadas as ações.¹²³⁹

Quando houver grande abrangência do suporte fático em sentido estrito (dados seus fatores de caracterização), poderá haver uma opção pela mais ampla restrição¹²⁴⁰ na individualização final, restrição esta que pode ser representada em termos de fatores próprios ou integrar a representação de fatores ou *ILC's*. A representação da aplicação hipotética de direito fundamental envolve, dessa maneira, um suporte fático estrito hipotético e restrições hipotéticas, pois, assim, garante-se, de um lado, a explicabilidade do sistema de IA judicial e, de outro, a possível derrotabilidade da solução hipotética mediante a correção argumentativa iniciada pelo aspecto humano do juiz ciborgue.

Todavia, o alcance prévio das restrições é uma escolha da equipe de desenvolvimento do sistema de IA judicial, na medida em podem ser estabelecidos uma concepção mais abrangente de suporte fático estrito e um alcance bastante

¹²³⁸ “Para que a consequência jurídica definitiva (proteção definitiva) de um direito fundamental ocorra, o suporte fático tem que ser preenchido e a cláusula de restrição, não; para que ela não ocorra é necessário ou que o suporte fático não seja preenchido ou que a cláusula de restrição o seja. Essa simples constatação é a base lógica para dois tipos muito distintos de construções fundamentadoras no âmbito dos direitos fundamentais: a da *teoria do suporte fático amplo* e a da *teoria do suporte fático restrito*” (ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 307).

¹²³⁹ “A partir da divisão entre o suporte fático e restrição, e essas duas respostas é possível atribuir quatro construções fundamentadoras, que podem ser representadas pelo seguinte esquema:

- (1) Suporte fático (+), restrição (+): não permitida.
- (2) Suporte fático (+), restrição (-): permitida.
- (3) Suporte fático (-), restrição (+): não permitida.
- (4) Suporte fático (-), restrição (-): não permitida.

Na primeira linha, a forma de agir (*h*) preenche o suporte fático (+) e é abrangida pela restrição (+) e, por isso, não é permitida; na segunda linha, *h* preenche o suporte fático e não é abrangida pela restrição (-) e, por isso, é permitida; etc. É perceptível que para a fundamentação de uma permissão apenas uma das construções serve como argumento, mas para a fundamentação de uma não-permissão, três. **Essas três construções são equivalentes no resultado.** Quem defende uma teoria *ampla* do suporte fático escolherá, se possível a primeira construção. Quem prefere uma teoria *restrita* irá se decidir, se possível, pela terceira fundamentação: o suporte fático não foi preenchido e, se o tivesse sido, a ação teria sido abrangida por uma restrição” (Ibid., p. 308).

¹²⁴⁰ “A frequente observação de que quanto mais amplo for o suporte fático mais ampla terá que ser a restrição diz respeito à relação entre a primeira e quarta construções” (Id.).

limitado das restrições¹²⁴¹, o que garante maior alcance de proteção ao direito fundamental, de maneira que as eventuais restrições mais limitantes do direito fundamental fiquem a cargo do intérprete humano ao se deparar com cada caso concreto. Na ausência de uma resposta padrão e totalmente fechada para esta questão (mediante uma teoria estrutural de suporte fático¹²⁴²), demonstra-se a adaptabilidade da teoria da argumentação e da teoria dos direitos fundamentais a cada contexto concreto de aplicação do Direito, abordagem que também pode ser vista na Ciência de Dados jurídica (vide o subcapítulo 5.6), o que denota uma ponte epistêmica entre as concepções interdisciplinares necessárias à via de comunicação inerente ao duplo canal de correção e ao paradigma de juiz ciborgue.

Frente às premissas até então referidas, uma regra só é uma “razão definitiva” no momento concreto de aplicação, pois é nesse momento em que é possível se verificar a ausência de casos de exceções trazidas pela derrotabilidade. A programação de sistemas de IA judicial pode se aproveitar disso, desde que respeite a autonomia do julgador. Contudo, caso se opte por programar os sistemas a partir dos princípios, o problema consiste em **quem irá individualizar princípios abstratos em regras definitivas para casos concretos**.

Esse problema ocorre porque, usualmente, quando um julgador infere uma regra a partir um princípio (ou da colisão de princípios) ele é a autoridade legítima para fazê-lo. No entanto, há a hipótese de o julgador responsável pela decisão não ter participado da fase de desenvolvimento do sistema de IA judicial, o que leva à discussão sobre a legitimidade de regras inferidas de princípios colocadas na programação. Nesse caso, a programação de regras inferidas a partir de princípios deve se basear em modelos já objetivados pelos **precedentes vinculantes**.

A partir disso, considerada a contextualização aqui exposta, percebe-se que será responsabilidade do julgador da causa escolher a opção que considera a juridicamente correta, dentre o rol de opções interpretativas sugeridas pela IA judicial.

¹²⁴¹ “Aquele que quiser proteger definitivamente o máximo possível dos direitos fundamentais poderá tentar fazê-lo por meio de uma concepção ampla de suporte fático e uma concepção restrita de restrição. Um suporte fático amplo não está necessariamente associado a uma restrição ampla. Decidir se ele deve estar associado a essa restrição ampla é uma questão normativa” (Id.).

¹²⁴² “Por isso, é necessário distinguir entre uma teoria ampla do suporte fático como uma *teoria estrutural* e a uma teoria ampla do suporte fático como uma *teoria normativa* orientada pelo resultado. Mas, na medida em que o caminho ou a construção de uma fundamentação podem ter influência no seu resultado, uma teoria ampla do suporte fático também tem, mesmo que na forma de uma teoria estrutural, um significado normativo” (Id.).

Essa escolha deve vir acompanhada das justificações argumentativas delineadas pelo julgador humano (ainda que com o auxílio do sistema de IA judicial). Além disso, o afastamento das demais opções trazidas pelo sistema de IA judicial (aquelas que não foram aceitas pelo julgador) deve ser justificado através de argumentação jurídica, com a exposição dos motivos jurídicos pelos quais não houve aceitação de cada uma das opções rejeitadas.

4.3.2 Elementos da Justificação Externa: Regras do Discurso Racional e sua Relação com a IA Judicial

O entrelaçamento entre sistemas baseados em Inteligência Artificial e justificação externa é necessário quando se pensa numa atuação pautada na pretensão de correção que nasce da relação comunicacional humano-máquina, uma vez que a justificação externa se volta à “fundamentação das premissas utilizadas na justificação interna”¹²⁴³, ainda que tais premissas já tenham sido testadas numa análise lógica.

Nesse enfoque, a justificação externa é composta por **regras de direito positivo**, a partir das quais se verificam os critérios de validade da norma perante o ordenamento jurídico.¹²⁴⁴ Por exemplo, a respeito da norma aplicada deve-se verificar se:

- a) há competência do ente federativo (União, Estado ou Município) para legislar na hipótese;
- b) adequa-se em relação à espécie legislativa na qual a norma é veiculada;¹²⁴⁵
- c) observa a constitucionalidade material (houve respeito aos direitos fundamentais, as demais normas do ordenamento constitucional foram observadas, etc.);
- d) respeita os direitos humanos¹²⁴⁶ internalizados no ordenamento jurídico brasileiro.

¹²⁴³ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 203.

¹²⁴⁴ Id.

¹²⁴⁵ Lei Complementar, Lei Ordinária, Medida Provisória ou outra espécie prevista no art. 59 da Constituição Federal.

¹²⁴⁶ “Desde a adesão do Brasil, sem qualquer reserva, ao Pacto Internacional dos Direitos Civis e Políticos (art. 11) e à Convenção Americana sobre Direitos Humanos - Pacto de San José da Costa Rica (art. 7º, 7), ambos no ano de 1992, não há mais base legal para prisão civil do depositário infiel, pois o caráter especial desses diplomas internacionais sobre direitos humanos lhes reserva lugar

Já nos **enunciados empíricos**¹²⁴⁷, a fundamentação se dá através da utilização de:

a) **métodos das ciências empíricas**, nos quais podem ser programados fatores que representem determinado suporte fático hipotético a respeito do domínio de certo ramo do conhecimento (por exemplo, verificada a ocorrência de uma determinada doença¹²⁴⁸ através de laudo médico aplica-se a isenção tributária correspondente¹²⁴⁹);

b) **máximas da presunção racional**;

c) **regras de ônus da prova no processo**.¹²⁵⁰

Na justificação externa, também são analisadas “**premissas que não são nem enunciados empíricos nem regras de direito positivo**”¹²⁵¹, o que se relaciona à argumentação jurídica em sentido estrito.

Analisando-se os três elementos do objeto da justificação externa, percebe-se que existem relações interativas entre os procedimentos, pois, para se aferir a validade de uma norma de direito positivo, muitas vezes, é necessário interpretar normas constitucionais, bem como para se aferir um fato empírico é necessário analisar regras de ônus da prova.¹²⁵²

Nesse sentido, Alexy classifica as formas de argumentos e as regras de justificação externa, de uma maneira ampla, em seis regras e formas sobre: **a)** interpretação; **b)** argumentação da Ciência do Direito (dogmática); **c)** uso de precedentes; **d)** argumentação prática geral; **e)** argumentação empírica; e **f)** formas especiais de argumentos jurídicos.¹²⁵³ Desse modo, a função da argumentação

específico no ordenamento jurídico, estando abaixo da Constituição, porém acima da legislação interna. O status normativo supralegal dos tratados internacionais de direitos humanos subscritos pelo Brasil torna inaplicável a legislação infraconstitucional com ele conflitante, seja ela anterior ou posterior ao ato de adesão” (BRASIL. Supremo Tribunal Federal (Plenário). **Recurso Extraordinário 4357**. Hierarquia Supralegal dos tratados e convenções sobre Direitos Humanos. Brasília, DF: Supremo Tribunal Federal, [2009]. Relator: Min. Carlos Ayres Britto. Relator para acórdão: Min. Gilmar Mendes, 03 de dezembro de 2008. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=AC&docID=595406>. Acesso em: 30 nov. 2022).

¹²⁴⁷ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 203.

¹²⁴⁸ A existência da doença está no domínio epistêmico da Medicina, não do Direito em si mesmo.

¹²⁴⁹ Por exemplo, a regra de isenção de imposto de renda prevista no art. 7º, XIV, da Lei 7.713/1998 (BRASIL. **Lei 7.713, de 22 de março de 1988**. Altera a legislação do imposto de renda e dá outras providências. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/l7713.htm. Acesso em: 30 nov. 2022).

¹²⁵⁰ Por exemplo, o art. 373 do Código de Processo Civil.

¹²⁵¹ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 203.

¹²⁵² Id.

¹²⁵³ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 204.

externa é possibilitar uma análise lógica das formas de argumentação, cujo resultado é o entendimento sobre a possibilidade e a necessidade de sua vinculação ao intérprete do Direito, especialmente através das interconexões trazidas pela argumentação prática geral e pela argumentação empírica.¹²⁵⁴

A partir dessas premissas, deve-se observar que “quase todas as formas de argumentação jurídica – igualmente quase todas as formas de argumentação prática geral – incluem enunciados empíricos”.¹²⁵⁵ A utilização da argumentação externa ganha contornos obrigatórios para o desenvolvimento de sistemas de Inteligência Artificial judicial que busquem, com graus adequados de determinabilidade e adaptabilidade em termos de programação (*factors* e *ILC's*) e modelagem de IA (*machine learning*), representar a plasticidade mundo jurídico com auxílio de **argumentação empírica** que atua na “justificação dos enunciados empíricos empregados na justificação interna”¹²⁵⁶ e “para a justificação de enunciados empíricos empregados na justificação externa de enunciados não-empíricos”.¹²⁵⁷

Para a fundamentação das decisões jurídicas, a argumentação empírica se volta aos problemas do conhecimento empírico, haja vista que os enunciados empíricos podem ser sobre: **a)** fatos singulares; **b)** ações concretas; **c)** motivos dos agentes; **d)** acontecimentos (presentes, passados e futuros); **e)** estado de coisas (presentes, passados e futuros); **f)** regularidade das ciências naturais ou das ciências sociais; **g)** a necessária cooperação interdisciplinar entre as diversas áreas de Ciência (Economia, Sociologia, Medicina, Psicologia, etc.).¹²⁵⁸ Esse conhecimento empírico pode ser tratado computacionalmente de maneira que seja relevante para a solução de casos (*case based reasoning*) ou para a interpretação de sistemas jurídicos (*statutory reasoning*) em domínios jurídicos bem delimitados¹²⁵⁹, de modo que a representação dos enunciados empíricos ocorra através de fatores (*factors*) ou

¹²⁵⁴ Ibid., p. 205.

¹²⁵⁵ Id.

¹²⁵⁶ Ibid., p. 228.

¹²⁵⁷ Id.

¹²⁵⁸ Id.

¹²⁵⁹ “En la teoría tradicional, y en las obras de filosofía jurídica, por ‘sistema jurídico’ o bien por ‘orden jurídico’ se entiende el derecho en su globalidad, la totalidad de las normas jurídicas. Pero, en la práctica, los jueces, abogados, juristas, etc., nunca están interesados en la totalidad del orden jurídico. Ellos siempre se ocupan de una selección (usualmente más bien pequeña) de disposiciones o normas jurídicas, que son relevantes para el problema que les interesa. En Normative Systems, por sistema jurídico se entiende un microsistema, no la totalidad del orden jurídico. Pienso que esta es una importante innovación. Sin embargo, muchos críticos de ese libro, con la notable excepción de Ricardo Guastini, aparentemente no advierten eso” (BULYGIN, Eugenio. **Lógica deóntica, normas y proposiciones normativas**, p. 53).

conceitos jurídicos intermediados (*ILC's*) moldados na programação de sistemas de IA judicial¹²⁶⁰, a fim de que o sistema de Inteligência Artificial possa, a partir dos enunciados empíricos, inferir consequências e possibilidades de aplicação do Direito, de maneira a auxiliar o julgador humano (agora o juiz ciborgue) na composição do seu esquema de argumentação externa.

Dessa maneira, a regra de transição **(6.1)** da teoria do discurso prático geral prevê que para “qualquer falante é, em qualquer momento, possível passar a um discurso teórico (empírico)”¹²⁶¹, o que também se aplica à justificação externa auxiliada por sistemas de Inteligência Artificial judicial.

Na fase da justificação externa, é necessário, também, se pensar nos “cânones de interpretação”¹²⁶², os quais se agrupam em seis grupos principais¹²⁶³: **a)** argumento semântico; **b)** argumento genético; **c)** argumento histórico; **d)** argumento comparativo; **e)** argumento sistemático; e **f)** argumento teleológico.¹²⁶⁴ Esses atuam para a fundamentação auxiliar à justificação interna¹²⁶⁵, de maneira que a estrutura lógica dos cânones de interpretação pode ser representada como uma decorrência da norma *R* quando se utiliza uma regra de uso das palavras *W*, do que se segue uma norma concreta *R'*. Ou seja, no aspecto lógico, um cânone de interpretação pode ser representado como uma variação linguística (*Mx*) da norma (conforme já descrito no subcapítulo 4.2), que se expressa da seguinte forma: **(2') (x) (Mx → ORx)**.¹²⁶⁶ Assim, a regra concreta surgida a partir do cânone de interpretação *R'* é uma interpretação da regra *R* através da regra de uso de palavras *W'*, ou seja *R_W*.¹²⁶⁷

¹²⁶⁰ “The particular kind of theory their model constructs, I maintain, however, does not correspond to a recognizably reasonable legal argument scheme. That would require a proposed test or rule for deciding a case, involving intermediate legal concepts, to play a central role. In addition, a more detailed representation of case facts will, I predict, be necessary for computationally modeling a contextually-sensitive comparison of cases to determine is an analogy drawn in terms of the concept applies or not” (ASHLEY, Kevin. Evaluating the Uses of Values in a Model of Legal Reasoning, p. 35-36).

¹²⁶¹ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 182.

¹²⁶² Ibid., p. 206.

¹²⁶³ A lista não é exaustiva, conforme aponta Alexy (Ibid., p. 206).

¹²⁶⁴ Ibid., p. 207.

¹²⁶⁵ Estrutura abreviada da justificação interna:

“(J.1.2').(1) (Tx → ORx) (R)
 (2) (Mx → Tx) (W)
 (3) Ma
 (4) ORa (1)-(2)”

(Ibid., p. 206).

¹²⁶⁶ Id.

¹²⁶⁷ Id.

Os cânones de interpretação têm por finalidade a fundamentação de: **a)** interpretações realizadas por meio de justificação interna (I^R_W); **b)** “normas não positivas”; e **c)** outros enunciados jurídicos.¹²⁶⁸

O **argumento semântico** se apresenta quando, no uso da linguagem, se justifica, se crítica ou se afirma como possível (admissível)¹²⁶⁹ uma interpretação **R'** de **R**, na qual **W** é uma “**especificação**”¹²⁷⁰ da linguagem natural ou da linguagem técnica (em particular a da Ciência do Direito – dogmática).¹²⁷¹ Nesse aspecto, percebe-se que o argumento semântico é campo atuação típico para a IA jurídica, baseada em representação de conhecimento (*factors* e *ILC's*) ou em processamento de linguagem natural, uma vez que as *especificações* representativas de significado podem ser programadas pelo assessor-desenvolvedor ou sugeridas por um sistema de processamento de linguagem natural.¹²⁷²

Assim, existem três formas argumentos semânticos que consistem em: **a)** “(J.3.1) **R'** deve ser aceita como interpretação de **R**, com base em W_i ”¹²⁷³; **b)** “(J.3.2) **R'** não pode ser aceita como interpretação de **R**, com base em W_k ”¹²⁷⁴; e **c)** “(J.3.3) É possível aceitar **R'** como interpretação de **R** e é possível não aceitar **R'** como interpretação de **R**, pois não regem nem W_i nem W_k ”.¹²⁷⁵ Desse modo, essa forma de utilização de argumentação em sistemas de Inteligência Artificial leva à afirmação da validade ou da não-validade de W_i e W_k . A entropia linguística a ser enfrentada pelo juiz ciborgue se apresenta no uso de argumentos semânticos, o que exige plasticidade da programação ou do modelo de IA, pois, diante da variabilidade inerente aos falantes (onde se inclui a fala auxiliada por IA) em relação à competência linguística, à investigação empírica, e à referência à autoridade de dicionários.¹²⁷⁶

Os argumentos **(J.3.1)** e **(J.3.2)** são definitivos caso se estabeleça que o caso **A** (representada pelo suporte fático hipotético ou pelo suporte fático concreto) se rege

¹²⁶⁸ Id.

¹²⁶⁹ ATIENZA, Manuel. Op. cit., p. 208.

¹²⁷⁰ Não há argumento semântico se **W** pressupõe uma “determinação” da decisão por uma das linguagens, pois tal “determinação” não se pode justificar por referência a um uso existente (são necessários outros argumentos) (ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 206-207).

¹²⁷¹ Ibid., p. 207.

¹²⁷² Especificamente em relação às técnicas de processamento de linguagem natural, deve se destacar que as mesmas são variadas, todavia o estado da arte da PLN (ou *NLP*) envolve a utilização de tecnologias *deep learning*.

¹²⁷³ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 207.

¹²⁷⁴ Id.

¹²⁷⁵ Id.

¹²⁷⁶ Id.

pela norma **R** (a interpretação **Mx** é permitida porque representa **R**) ou que isso não ocorre (a interpretação **Mx** é proibida porque não representa **R**).

No mais, a utilização do argumento **(J.3.3)** demonstra que apenas com meios semânticos não se pode alcançar uma decisão (**T** é vago e **A** entra no campo de vagueza de **T**), o que, mais que uma “especificação” da linguagem, demanda uma “determinação”, por meio de outros tipos de argumentos.¹²⁷⁷

O argumento genético se pauta na justificação de “uma interpretação de **R'** em **R**, porque corresponde à vontade do legislador”.¹²⁷⁸ Existem duas formas de argumentos genéticos¹²⁷⁹, sendo que na primeira forma **(J.4.1)** a interpretação (I^R_W) = **R'** é o objeto da vontade¹²⁸⁰ do legislador¹²⁸¹. Já a segunda forma **(J.4.2)** é uma variante do argumento teleológico, quando se afirma que o legislador perseguiu¹²⁸² com **R**¹²⁸³ os fins **Z₁, Z₂... Z_n** na combinação com **K**, sendo que $([Z_1, Z_2... Z_n]K) \leftrightarrow Z$.¹²⁸⁴

Inicialmente, observa-se que o argumento da vontade do legislador se baseia numa forma relativamente fraca do argumento genético, pois nesse caso “a vontade do legislador é só uma razão para uma interpretação”¹²⁸⁵, o que abre a possibilidade para razões contrárias na programação de IA judicial.

Por outro lado, a forma forte da regra de inferência presente em **(J.4.2)**¹²⁸⁶ se baseia em um esquema¹²⁸⁷ de uma variante do silogismo prático **(S)** na qual é obrigatório alcançar um estado de coisas **Z**, de modo que, se não ocorrer **M**, não se

¹²⁷⁷ Ibid., p. 208.

¹²⁷⁸ Id.

¹²⁷⁹ ATIENZA, Manuel. Op. cit., p. 208.

¹²⁸⁰ “(a) É uma razão para a validade de **R'** que o legislador deseje que **R** se interprete mediante **W** ($I^R_W + R'$)” (ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 209).

¹²⁸¹ (J.4.1) .(1) **R'** ($=I^R_W$) é querido pelo legislador.
.(2) **R'**

¹²⁸² “(b) Que o legislador persiga com **R** o fim **Z** é uma razão para que na aplicação de **R** seja obrigatório perseguir **Z**.

(c) Se é obrigatório perseguir o fim **Z**, então também é obrigatório qualquer meio necessário para a realização de **Z**” (Id.).

¹²⁸³ A validade de **R** na interpretação I^R_W é necessária para a realização de $(Z_1, Z_2... Z_n)K$.

¹²⁸⁴ (J.4.2) .(1) Com **R** o legislador pretende alcançar **Z**.

.(2) $\neg R' (=I^R_W) \rightarrow \neg Z$

.(3) **R'**

¹²⁸⁵ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 209.

¹²⁸⁶ Regra de inferência (C).

¹²⁸⁷ (S) .(1) **OZ**

.(2) $\neg M \rightarrow \neg Z$

.(3) **OM**

alcança **Z**, ou seja, **M** é uma condição para a ocorrência de **Z**¹²⁸⁸. Para verificar a validade desta relação argumentativa (**S**), seriam necessárias considerações em relação à lógica proposições normativas¹²⁸⁹ e em relação à teoria da relação meio-fim.

Todavia, a utilização de argumento genético exige a observância do **requisito de saturação**¹²⁹⁰, pois é incompleta a argumentação que se baseia na simples utilização da afirmação de que determinada interpretação expressa a vontade do legislador ou o fim da norma, eis que, para auxiliar o argumento genético (de maneira que não se resuma a um simples argumento de autoridade), são necessários outros tipos de argumentos aqui estudados.¹²⁹¹ Dessa maneira, o que se busca é o estabelecimento de um suporte fático abstrato de maneira que se tenha ciência das dificuldades do uso dos argumentos genéticos, pois “não é claro quem se deve considerar como sujeito da ‘vontade do legislador’”¹²⁹², uma vez que, frequentemente, não é possível determinar o conteúdo dessa vontade.

Por sua vez, o **argumento histórico** se caracteriza quando se interpreta o suporte fático concreto (os fatos) relacionado à historicidade do problema jurídico em análise, de maneira que os fatores podem ser razões contrárias ou favoráveis à interpretação.¹²⁹³ Ou seja, o argumento trata do aprendizado da história da interpretação do Direito, a partir da análise de consequências históricas, sociológicas e econômicas. Desse modo, o argumento histórico se caracteriza¹²⁹⁴ pelo fato de que já houve no passado uma solução para o problema representado no sistema de IA judicial, a qual levou à consequência indesejável **F**.¹²⁹⁵ Assim, se a consequência **F** é indesejável, ela é proibida em termos de lógica deôntica¹²⁹⁶ para novas situações semelhantes a serem enfrentadas pelo sistema de IA judicial.

¹²⁸⁸ “(1) É obrigatório alcançar o estado de coisas **Z**.”

(2) Se não ocorrer **M**, então não se alcança **Z** (isto é, **M** é uma condição para **Z**).

(3) É obrigatório que ocorra **M**” (ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 209).

¹²⁸⁹ BULYGIN, Eugenio. **Lógica Deôntica, Normas y Propositiones Normativas**, p. 52.

¹²⁹⁰ ATIENZA, Manuel. Op. cit., p. 209.

¹²⁹¹ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 210.

¹²⁹² Id.

¹²⁹³ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 211.

¹²⁹⁴ “(1) Já se deu uma vez determinada solução para o problema discutido; (2) esta conduziu à consequência **F**; (3) **F** é indesejável; (4) as situações não são entre si diferentes para que **F** não ocorra hoje; (5) portanto, a solução em questão não é aceitável hoje” (Id.).

¹²⁹⁵ Id.

¹²⁹⁶ ALCHOURRÓN, Carlos E.; BULYGIN, Eugenio. Op. cit., p. 10.

Logo, o que se tem no caso do argumento histórico, é uma programação jurídica da não aceitação da consequência (que está no âmbito do suporte fático concreto) em razão de sua já demonstrada inadequação ao Direito, o que impede que o sistema de IA judicial indique determinada a solução interpretativa como possível ou que acuse a consequência indesejada, de maneira que o julgador humano, caso venha a adotar a interpretação diversa da programada ou modelada, tenha que obrigatoriamente elaborar uma justificação superatória do argumento histórico que lhe seja contrário. Destaca-se que o argumento histórico sempre deve ter uma **premissa normativa** que deve ser fundamentada¹²⁹⁷, pois, do contrário, haveria um mero consequencialismo que não se afigura legítimo no recorte jurídico-epistêmico adotado no presente estudo.

Também devem ser destacados os **argumentos comparativos**, nos quais são tomados como referência suportes fáticos que representam o “estado de coisas jurídicas de outra sociedade”¹²⁹⁸, que remetem ao estudo do Direito Comparado e sua necessária intermediação intersubjetiva e histórica, o que também deve incluir na programação da IA judicial pelo menos uma premissa normativa vigente¹²⁹⁹ juntamente com as premissas empíricas (presentes nos *factors* ou nos *ILC's*), típicas do suporte fático aplicável. Dessa forma, a implementação de argumento comparativo pode ser realizada de maneira semelhante, mas com adaptações, em relação àquela feita no argumento histórico¹³⁰⁰, mediante a proibição ou a não recomendação da consequência jurídica indesejável.

O **argumento sistemático** leva em conta principalmente a “relação lógica ou teleológica de uma norma com outras normas, fins e princípios”¹³⁰¹, o caracterizando como um argumento sistemático-teleológico¹³⁰², uma vez que no “sentido estrito, só devem considerar-se como argumentos sistemáticos aqueles que se referem unicamente às relações lógicas entre as normas”.¹³⁰³ Desse modo, a demonstração de **contradição normativa** é a forma mais importante de argumento sistemático: a) se **R’₁** (proposição normativa de **R₁** por meio de **W**) contradiz a proposição normativa

¹²⁹⁷ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 211.

¹²⁹⁸ Id.

¹²⁹⁹ Ainda que se trata da utilização de direito comparado, é necessário que se esteja a aplicar normas constitucionais e infraconstitucionais vigentes no território brasileiro.

¹³⁰⁰ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 211.

¹³⁰¹ Id.

¹³⁰² Id.

¹³⁰³ Id.

R₂ que se deve reconhecer como válida, então se deve abandonar **R₁**; b) a contradição somente se configura se “se todas as interpretações de **R₁** forem excluídas por outras normas”¹³⁰⁴, o que é raro mas pode ocorrer. Nesse aspecto, deve ser destacado que a relação sistemático-teleológica se dá entre proposições normativas diversas, oriundas de um mesmo sistema normativo¹³⁰⁵, o que possibilita o desenvolvimento de sistemas de IA judicial que se pautem em coerência interna e consistência, de maneira a auxiliar o julgador humano.

Já o **argumento teleológico**¹³⁰⁶ se pauta nos fins racionais objetivamente aferíveis a partir do sistema jurídico, o que ocorre para além da análise teleológica que se dá a partir da inferência de fins buscados pelo legislador histórico, pois não são buscados os fins subjetivos de pessoas do presente ou do passado.¹³⁰⁷ Assim, se criam os chamados argumentos teleológico-objetivos¹³⁰⁸, pois nessa espécie de argumento teleológico os fins racionais se baseiam na concepção de “sujeito hipotético racional”¹³⁰⁹ (ou seja, “a comunidade de quem deve tomar decisões”¹³¹⁰) como quem estabelece a argumentação necessária para que ocorram decisões a partir do sistema jurídico, de modo que as proposições normativas finalísticas elaboradas pelo intérprete humano-máquina “são hipóteses sobre os fins estabelecidos por este sujeito hipotético”¹³¹¹.

Acerca do **argumento teleológico** deve ser observado que em sua interação com a os valores presentes em princípios¹³¹², há um aprofundamento do duplo canal de correção típico do juiz ciborgue para o campo da pretensão de correção moral, sem abrir mão da pretensão de correção normativa, já que a representação de posições axiológicas é inerente à aplicação de princípios (o que mantém a argumentação racional nos limites do sistema jurídico) e conduz a pretensão de correção moral para dentro da programação ou da modelagem do sistema de IA judicial.

¹³⁰⁴ Id.

¹³⁰⁵ BULYGIN, Eugenio. **Lógica deóntica, normas y proposiciones normativas**, p. 53.

¹³⁰⁶ “Um esclarecimento completo do conceito de *argumento teleológico* pressupõe uma análise detalhada dos conceitos de fim e de meio, assim como dos conceitos vinculados aos de vontade, intenção, necessidade prática e fim” (ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 211).

¹³⁰⁷ Id.

¹³⁰⁸ Ibid., p. 212.

¹³⁰⁹ Id.

¹³¹⁰ Id.

¹³¹¹ Id.

¹³¹² ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 147-153.

Além disso, de maneira a se superar o empirismo, os fins racionais da interpretação teleológica caracterizam-se, normativamente, como **um estado de coisas ou um fato prescrito** (um suporte fático), o que pode ser simbolizado por **OZ**. A necessidade de interpretação $I^R_w = R'$ para alcançar o fim **Z** (sua representação com um meio para **Z**) pode ser assim indicada: $\neg R' (=I^R_w) \rightarrow Z$.¹³¹³ Dessa forma, a regra argumentativa **(J.5)**¹³¹⁴ é a forma mais simplificada da expressão do argumento teleológico:

- (J.5) .(1) OZ
 .(2) $\neg R' (=I^R_w) \rightarrow \neg Z$
 .(3) R'

Destaca-se que **(J.5)** se relaciona com a já referida forma de argumento genético **(J.4.2)**, com a diferença de que em **(J.5)** não é buscado o fim **Z** como algo desejado pelo legislador, haja vista que o fim **Z** é objetivado a partir do sistema jurídico.

Por outro lado, uma visão geral do argumento teleológico em que **(J.5)** e **(J.4.2)** atuam conjuntamente é a seguinte:

- (S) .(1) OZ
 .(2) $\neg M \rightarrow \neg Z$
 .(3) OM

Dessa forma, em (1) está contido o enunciado normativo **(J.5)**, o qual pressupõe a validade de duas premissas complementemente diferentes “**OZ**”; em (2) há a representação do enunciado empírico “ $\neg M \rightarrow \neg Z$ ”, o qual para ser fundamentado necessita do conhecimento das “legalidades empíricas”¹³¹⁵ que remetem à utilização de argumentos empíricos.¹³¹⁶

Por sua vez, **(1)** é quase sempre fundamentado conforme a norma a ser aplicada. Ou seja, desde que a proposição normativa **R** seja válida, a partir do

¹³¹³ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 212.

¹³¹⁴ Ibid., p. 213.

¹³¹⁵ Id.

¹³¹⁶ Id.

argumento teleológico está prescrita a realização do estado de coisas **Z**.¹³¹⁷ Todavia, deve ser destacado que o estado de coisas **Z** frequentemente se fundamenta não apenas em uma única norma, mas em um **grupo de normas**.¹³¹⁸ Contudo, se **Z** não se segue logicamente da(s) norma(s) que são utilizadas para sua fundamentação em **(1)**, serão necessários novos argumentos auxiliares, salvo se uma das normas prescrever diretamente a realização de **Z**.¹³¹⁹

Para além de uma análise superficial do argumento teleológico, deve se destacar que a uma norma (ou a um grupo de normas) podem ser atribuídos **diversos fins que se excluem entre si** ou que, na prática, são realizados mediante sua **limitação recíproca**.¹³²⁰ Dessa maneira, quando se verifica uma determinada combinação de fins **([Z₁, Z₂, ..., Z_n] K)**¹³²¹, a programação dos *factors* necessários à implementação de argumento teleológico deve pressupor a aplicação de regras de preferência entre os fins representados ou, por outro lado, deve levar a uma representação que já considere previamente as preferências inseridas em um conceito jurídico intermediado (*ILC*).

No mais, os problemas mais difíceis da argumentação teleológica ocorrem quando **Z** ou **([Z₁, Z₂, ..., Z_n] K)** não podem ser descritos somente por meio de expressões empíricas, o que ocorre quando **Z é um estado de coisas** (presente ou futuro) definido de maneira que nele regem determinadas normas.¹³²² Ou seja, **Z** pode ser determinado por um estado de coisas (um suporte fático) regido pela norma fundamentada em **R**, nesse caso, a referência a **Z** só cumpre a função de esclarecer o que significa a validade da proposição normativa **R'**. Contudo, ainda que para a descrição de tal estado de coisas sejam necessárias proposições normativas com grande generalidade (normas de tipo geral ou princípios)¹³²³, a programação e a modelagem de sistemas de IA judicial devem levar à criação de soluções de auxílio ao julgador humano. Nessas hipóteses de disputa entre diversas finalidades a serem aplicadas de maneira simultânea num mesmo caso, o canal de correção humano deve ser acionado para a mediação da entropia interpretativa natural ao problema. Logo, nessa hipótese, **Z** é um estado de coisas regido pelos diversos princípios **(P₁, P₂, ...,**

¹³¹⁷ Id.

¹³¹⁸ Id.

¹³¹⁹ Id.

¹³²⁰ Ibid., p. 213-214.

¹³²¹ Ibid., p. 214.

¹³²² Id.

¹³²³ Id.

P_n), cada qual com o seu próprio suporte fático abstrato, o que demandará bastante plasticidade e adaptabilidade do assessor-desenvolvedor ao moldar o sistema de IA judicial e ao julgador humano, uma vez que é necessária uma concreção dos princípios com base em novas proposições normativas.

Assim, o que se demonstra a partir do argumento teleológico (e da interpretação de princípios) é que, definitivamente, não existe milagre e a IA não é panaceia, pois a inclusão do canal de correção advindo da Inteligência Artificial passa a exigir mais esforço justificatório no canal de correção humano. Logo, o juiz ciborgue é mais forte, não porque seja rápido, porque seja perfeito ou porque busque resolver problemas impossíveis¹³²⁴, mas porque **argumenta melhor**¹³²⁵ (com mais dados, com melhor percepção do suporte fático concreto e com melhor representação do suporte fático abstrato) na via comunicacional do duplo canal de correção.

Em relação às funções dos tipos de argumentos aqui estudados (o que Alexy chama de “cânones do discurso jurídico”¹³²⁶), seis pontos devem ser distinguidos a seguir.

O primeiro deles é o a) **campo de sua aplicabilidade**, no qual, excetuada a interpretação semântica, as demais formas de argumentos podem também se utilizar de uma multiplicidade de novos contextos (conflito de normas, restrição do campo de aplicação das normas e fundamentação de normas que não se podem deduzir da lei).

Também deve ser analisado o b) **status lógico** de cada espécie de argumento, na medida em que argumentos não são regras absolutas, são “simples pontos de vista ou linhas de investigação”¹³²⁷, uma vez que caracterizam a estrutura do ordenamento jurídico por meio de uma gramática da argumentação jurídica, mostrando “como se deve proceder para alcançar um resultado racional”.¹³²⁸ Nesse ponto, destaca-se que o *status* lógico dos diversos tipos de argumentos se mostra um objeto adequado para o desenvolvimento de sistemas de Inteligência Artificial judicial, sendo que o risco de

¹³²⁴ A mitologia grega já foi bastante comentada durante esta obra, mas lembre-se que a referência é ao juiz *Hércules*, caso a esta altura do texto ainda exista algum leitor crédulo em semideuses.

¹³²⁵ “Eu tentei, em um outro lugar, formular o sistema de regras do discurso com o auxílio de 28 regras tão completamente quanto possível. As regras estendem-se de tais, que exigem a liberdade da contradição, clareza idiomática, verdade empírica e sinceridade, sobre tais, que dão expressão, entre outras coisas, à ideia de generalizabilidade pelo fato de assegurarem o direito de participação de cada um em discurso e a consideração igual de cada um no discurso, até a tais, que são dirigidas à argumentação das consequências, à ponderação e à análise do nascimento de convicções normativas” (ALEXY, Robert. **Direito, Razão, Discurso**, p. 89).

¹³²⁶ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 214.

¹³²⁷ *Ibid.*, p. 215.

¹³²⁸ *Ibid.*, p. 216.

decisões meramente mecânicas é enfrentado pelo necessário canal de correção humano.

Por sua vez, no c) **requisito de saturação** um “argumento de uma forma só é completo se contém todas as premissas pertencentes a esta forma”¹³²⁹, já que a exigência de saturação “exclui a simples afirmação de que um argumento pode ser resultado de uma determinada interpretação literal, histórica ou finalista da norma”.¹³³⁰ Assim, a regra argumentativa **(J.6)** define que deve “ser saturada toda forma de argumento que houve entre os cânones da interpretação”.¹³³¹ Desse modo, uma interpretação jurídica deve passar por um rol de testes lógico-argumentativos em cada tipo de argumento disponível, de maneira a ser observado o requisito de saturação. O duplo canal de correção é um catalizador dessa tarefa, pois o sistema de IA judicial pode tentar uma série de testes pré-programados ou modelados por meio de *machine learning*, de maneira que os testes possam ser analisados pelo interprete humano e reprocessados pelo sistema de IA. Em suma, o juiz ciborgue sempre deve atuar buscando cumprir o seu dever de saturação.

Destaca-se que cada uma das **formas de argumentos** tem uma função que lhe é proeminente, contudo a partir dessas funções, não é possível se estabelecer uma hierarquia entre eles, ainda que haja a discussão entre a teoria objetiva e a teoria subjetiva da interpretação (vontade do legislador, argumentos teleológicos objetivos e argumentos sistemáticos).¹³³²

Nesse caminho, acerca do **problema da hierarquia** entre os tipos de argumentos no Direito, a teoria do discurso jurídico considera, também, a dimensão pragmática da fundamentação, sendo que não pretende proporcionar critérios tão fortes para a determinação com segurança de um resultado, aumentando a probabilidade de uma conclusão correta (racional).¹³³³ Desse modo, como um critério inicial em qualquer argumentação jurídica, Alexy propõe a regra argumentativa **(J.7)**:

“(J.7) Os argumentos que expressam uma vinculação ao teor literal da lei ou à vontade do legislador histórico prevalecem sobre os outros argumentos, a

¹³²⁹ Id.

¹³³⁰ Id.

¹³³¹ Id.

¹³³² Ibid., p. 217.

¹³³³ Ibid., p. 218.

não ser que se possam apresentar motivos racionais que deem prioridade a outros argumentos”.¹³³⁴

Contudo, a prevalência inicial de argumentos pautados na vinculação ao teor literal da lei e à vontade do legislador histórico, embora seja uma deferência constitucional ao Poder Legislativo (art. 5º, II, da Constituição Federal), pode ser superada por meio de justificativas racionais de por que é atribuído maior peso a uma outra determinada espécie de argumento, o que passa pelo dever de saturação e pela justificação racional inerente à pretensão de correção do discurso jurídico, a fim de afastar o mero arbítrio.¹³³⁵ Dessa forma, as “determinações de peso aferidas pelos participantes no discurso nos diferentes contextos de interpretação valem, porque devem ser obtidas razões, não só para o caso discutido, mas para todos os casos iguais nos aspectos relevantes”.¹³³⁶ De maneira a sintetizar esta concepção como uma variante do princípio da universalidade, a regra argumentativa **(J.8)** prevê que a “determinação do peso de argumentos de diferentes formas deve ocorrer segundo regras de ponderação”¹³³⁷. Todavia, deve se ter em vista que as regras de ponderação variam conforme cada ramo do Direito, o que inibe uma certeza teórica definitiva, inserindo o problema no campo da necessária pluralidade epistêmica e sua relação com a argumentação dogmática (vide o subcapítulo 3.3), o que representa um desafio ao assessor-desenvolvedor na busca de moldar os sistemas de IA judicial de maneira a garantir o pleno exercício da legitimidade discursiva no processo de argumentação jurídica, pois a regra argumentativa **(J.9)** determina que se deve “levar em consideração todos os argumentos possíveis e que possam ser incluídos por sua forma entre os cânones da interpretação”.¹³³⁸

Em relação à argumentação dogmática típica da Ciência do Direito¹³³⁹, o que inclui as regras argumentativas **(J.10)**¹³⁴⁰, **(J.11)**¹³⁴¹ e **(J.12)**¹³⁴², rememora-se que o

¹³³⁴ Id.

¹³³⁵ Id.

¹³³⁶ Id.

¹³³⁷ Ibid., p. 219.

¹³³⁸ Id.

¹³³⁹ Ibid., p. 219-237.

¹³⁴⁰ “(J.10) Todo enunciado dogmático, se é posto em dúvida, deve ser fundamentado mediante o emprego, pelo menos, de um argumento prático de tipo geral” (Ibid., p. 259).

¹³⁴¹ “(J.11) Todo enunciado dogmático deve enfrentar uma comprovação sistemática, tanto em sentido estrito como em sentido amplo” (Id.).

¹³⁴² “(J.12) Se são possíveis argumentos dogmáticos, devem ser usados” (Id.).

tema¹³⁴³ foi trabalhado no contexto da pluralidade epistêmica necessária para desenvolvimento de sistemas de IA judicial, remetendo o debate para o subcapítulo 3.3 desta obra, sendo desnecessária a repetição do tema nesse ponto.

Ainda no âmbito da justificação externa, ganha destaque a utilização de **precedentes**, o que se faz necessário no contexto do Direito brasileiro devido à força vinculativa que os precedentes passaram a ter nos últimos tempos, em especial, após o advento do Código de Processo Civil de 2015.¹³⁴⁴ Em relação a esse tema, não é objeto desta obra realizar um estudo sistemático sobre a teoria dos precedentes judiciais e suas possibilidades mediante a implementação de sistemas computacionais dedicados à tarefa de análise de precedentes (o que seria um estudo à parte), sendo que o destaque se relaciona à sua importância para a teoria da argumentação no contexto da presente pesquisa.

Feita a necessária ressalva acima, deve-se observar que os precedentes podem ser observados a partir da teoria do discurso jurídico, especialmente no que se refere à relação de argumentos baseados em precedentes com outros argumentos possíveis de utilização no discurso jurídico.¹³⁴⁵

A partir dessa premissa, evidencia-se que o uso de precedentes tem fundamento no princípio da universalidade, haja vista o fato de que “toda decisão formula uma pretensão de correção”¹³⁴⁶, o que leva uma **carga argumentativa** a quem pretenda se afastar de um precedente mediante razões suficientes para tanto.¹³⁴⁷

Destaca-se que a regra da carga de argumentação, ao demandar esforço do juiz ciborgue através do duplo canal de correção, apresenta funções de estabilização¹³⁴⁸, progresso¹³⁴⁹ e descarga¹³⁵⁰, o que aproxima os precedentes dos enunciados jurídicos estudados no âmbito da argumentação dogmática¹³⁵¹, uma vez que os precedentes podem ser representados por proposições normativas que, no nível da programação e da modelagem de sistemas de IA judicial, serão

¹³⁴³ ATIENZA, Manuel. Op. cit., p. 211.

¹³⁴⁴ BRASIL. **Lei 13.105, de 16 de março de 2015.**

¹³⁴⁵ ALEXYY, Robert. **Teoria da Argumentação Jurídica**, p. 237.

¹³⁴⁶ Ibid., p. 238.

¹³⁴⁷ Id.

¹³⁴⁸ A função de estabilização se volta à fixação dos enunciados jurídicos (proposições normativas), os quais se fazem reproduzíveis em novos casos.

¹³⁴⁹ A função de progresso que possibilita a evolução da interpretação do Direito.

¹³⁵⁰ A função de descarga é voltada adoção, nas fundamentações, de enunciados já comprovados e aceitos, pelo menos de maneira provisória.

¹³⁵¹ ALEXYY, Robert. **Teoria da Argumentação Jurídica**, p. 239.

representados através de fatores (*factors*) ou conceitos jurídicos intermediados (*ILC's*). Desse modo, o uso dos precedentes na argumentação jurídica voltada à implementação de sistemas de Inteligência Artificial judicial representa uma deferência à segurança jurídica¹³⁵² e à proteção da confiança¹³⁵³ necessárias ao duplo canal de correção.

Dessa forma, o respeito ao princípio da universalidade inibe a existência de soluções incompatíveis entre si (incompatibilidade esta que aumenta o grau de inconsistência dos sistemas jurídicos¹³⁵⁴), o que “não exclui a possibilidade de não se seguir o precedente cuja máxima de decisão seja reconhecida como equivocada”.¹³⁵⁵ A partir disso, o fomento à ideia de um juiz ciborgue que observe os precedentes¹³⁵⁶ passa pelo respeito às regras argumentativas de que quando “se puder citar um precedente a favor ou contra uma decisão, deve-se fazê-lo” (**J.13**), bem como que quem “quiser se afastar de precedente, assume a carga da argumentação” (**J.14**).¹³⁵⁷

Nesse sentido, a regra do discurso jurídico (**J.14**) exige o cumprimento do dever de justificação inerente ao duplo canal de correção, o que passa pela definição do que se qualifica como *ratio decidendi*¹³⁵⁸ de cada precedente (a proposição normativa a ser representada no sistema de IA judicial), de maneira a diferenciá-lo do que foi tratado em mero caráter de *obiter dictum*¹³⁵⁹ nos julgados.

¹³⁵² “Quando a doutrina e os Tribunais brasileiros fecham os olhos para esta realidade, de que a Constituição é mais do que seu texto e que sua estrutura calcada em direitos e princípios fundamentais de caráter aberto e abstrato aceitam a densificação de seu conteúdo de acordo com os valores culturais, históricos da sociedade, acaba por fingir que o texto constitucional permite a segurança almejada por seus cidadãos, consentindo, ao contrário, que o Judiciário fique livre para decidir por pressões políticas, pragmáticas e contingenciais, enfraquecendo a própria noção da Constituição” (BARBOZA, Estefânia Maria de Queiroz. Op. cit., p. 157).

¹³⁵³ MARINONI, Luiz Guilherme. **Precedentes Obrigatórios**. 2ª ed. São Paulo: Revista dos Tribunais 2011, p. 107-109.

¹³⁵⁴ ALCHOURRÓN, Carlos E.; BULYGIN, Eugenio. Op. cit., p. 17.

¹³⁵⁵ ALEX, Robert. **Teoria da Argumentação Jurídica**, p. 240.

¹³⁵⁶ ATIENZA, Manuel. Op. cit., p. 211.

¹³⁵⁷ ALEX, Robert. **Teoria da Argumentação Jurídica**, p. 240.

¹³⁵⁸ “Um fundamento ou motivo, embora não necessário, pode ser suficiente para se alcançar a decisão. O motivo suficiente, porém, torna-se determinante apenas quando, individualizado na fundamentação, mostra-se como premissa sem a qual não se chegaria à específica decisão. Motivo determinante, assim, é o motivo que, considerado na fundamentação, mostra-se imprescindível à decisão que foi tomada. Este motivo, por imprescindível, é essencial, ou melhor, é determinante da decisão. Constitui a *ratio decidendi*” (MARINONI, Luiz Guilherme. Op. cit., p. 293).

¹³⁵⁹ “Nos sistemas de tradição de *common law* é necessário verificar que nem toda a opinião manifestada pelo juiz cria um precedente judicial. Para que uma opinião tenha o peso de um precedente é necessário que ela tenha sido dada por um juiz e também que tenha sido necessária para a decisão de um caso particular, isto é, não pode ser *obiter dictum*. A *ratio decidendi* de uma decisão nos sistemas de *common law* é o que vai criar um precedente vinculante nas decisões seguintes e, portanto, se faz necessário descobri-la” (BARBOZA, Estefânia Maria de Queiroz. Op. cit., p. 184-185).

Neste ponto, deve ser destacado que muitos projetos de IA judicial no Brasil têm foco na análise de precedentes (o maior exemplo é o *Projeto Victor* do Supremo Tribunal Federal¹³⁶⁰). O campo de desenvolvimento de sistema de IA judicial é bastante fértil e encontra grande potencialidade de aplicabilidade prática nos mais variados campos do Direito, desde a criação de sistemas de recuperação de precedentes até o uso de processamento de linguagem natural¹³⁶¹ para o auxílio na tarefa de delineamento da *ratio decidendi*.¹³⁶²

Sendo assim, com o auxílio de sistemas de IA judicial na via comunicativa humano-máquina, a definição da *ratio decidendi* dos precedentes é um passo necessário para a representação de proposições normativas e para a realização dos procedimentos argumentativo-institucionais de *distinguishing* e *overruling*¹³⁶³ que são exigidos pelo art. 489, §1º, V e VI, do Código de Processo Civil.¹³⁶⁴

No mais, existem **formas de argumentos jurídicos especiais**¹³⁶⁵ que também podem ser representadas através de lógica proposicional, de maneira que existe a potencialidade de serem incorporadas em sistemas de IA judicial. Alexy exemplifica os argumentos jurídicos especiais no *argumentum a contrario*, na analogia e no *argumentum ad absurdum*, os quais são melhor compreendidos a partir da sua expressão lógica.¹³⁶⁶

O a) ***argumentum a contrario*** atua de maneira a fechar a possibilidade de interpretações que não sejam expressamente previstas ou autorizadas a partir das

¹³⁶⁰ HARTMANN PEIXOTO, Fabiano. Projeto Victor: Relato do Desenvolvimento da Inteligência Artificial na Repercussão Geral do Supremo Tribunal Federal.

¹³⁶¹ BONAT, Debora; HARTMANN PEIXOTO, Fabiano. **Racionalidade no Direito: Inteligência Artificial e Precedentes**. Curitiba: Alteridade, 2020, p. 117-140.

¹³⁶² PEREIRA, João Sérgio dos Santos Soares. Op. cit., p. 426.

¹³⁶³ “Entretanto, deve-se ter em vista que a doutrina do *stare decisis* não significa apenas a vinculação do juiz aos precedentes no sentido de seguir o precedente, mas também inclui as possibilidades de superação e revogação (*overruling*), bem como de distinção dos precedentes (*distinguishing*), na medida em que em todos estes casos há compromisso com o passado” (BARBOZA, Estefânia Maria de Queiroz. Op. cit., p. 182).

¹³⁶⁴ “Art. 489. São elementos essenciais da sentença:

(...)

§ 1º Não se considera fundamentada qualquer decisão judicial, seja ela interlocutória, sentença ou acórdão, que:

(...)

V - se limitar a invocar precedente ou enunciado de súmula, sem identificar seus fundamentos determinantes nem demonstrar que o caso sob julgamento se ajusta àqueles fundamentos;

VI - deixar de seguir enunciado de súmula, jurisprudência ou precedente invocado pela parte, sem demonstrar a existência de distinção no caso em julgamento ou a superação do entendimento” (BRASIL. **Lei 13.105, de 16 de março de 2015**).

¹³⁶⁵ ATIENZA, Manuel. Op. cit., p. 212.

¹³⁶⁶ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 241.

normas, pois veda que sejam admitidas consequências jurídicas não previstas explicitamente (o que leva a um “*argument from a distinction due to a missing feature*”¹³⁶⁷). Caso as consequências não sejam verificadas, não é caso de aplicação da proposição normativa. Sua função lógica na teoria da argumentação é uma variação da regra “se, então” (“*if, then*”), típica da programação computacional, o que denota sua aplicabilidade prática no desenvolvimento de IA judicial. A expressão lógica do *argumentum a contrario* é sintetizada por Alexy¹³⁶⁸ na regra argumentativa **(J.15)**:

$$\begin{aligned} \text{(J.15)} \quad &.(1) (x) (OGx \rightarrow Fx) \\ &.(2) (x) (\neg Fx \rightarrow \neg OGx) \quad (1) \end{aligned}$$

Já a b) **analogia**, que consiste na aplicação de uma solução jurídica, existente para uma determinada hipótese (proposição normativa), para uma outra hipótese que guarda similaridade (se caracteriza como um “*argument from sufficient similarity*”¹³⁶⁹) em relação ao suporte fático, mas que não possui uma solução explícita no sistema jurídico. Nesse sentido, Kevin Ashley destaca que na analogia há o compartilhamento de fatos subsumidos (suporte fático) a conceitos jurídicos intermediados¹³⁷⁰, o que possibilita a programação¹³⁷¹ de sistemas de IA com a utilização deste tipo de argumento jurídico especial. Assim, a analogia pode ser formalizada através da regra argumentativa **(J.16)** proposta por Alexy¹³⁷²:

$$\text{(J.16)} \quad .(1) (x) (Fx \vee F \text{ sim } x \rightarrow OGx)^{1373}$$

¹³⁶⁷ ASHLEY, Kevin. Evaluating the Uses of Values in a Model of Legal Reasoning, p. 32.

¹³⁶⁸ “Deve-se estabelecer que só se x é F , deve-se seguir a consequência jurídica em questão. Que isso ocorra ou não no caso depende de como se interpreta. Pode-se com certeza apoiar-se muitas vezes no texto das normas; porém frequentemente são necessários argumentos adicionais para tanto. São admissíveis como argumentos adicionais todos os argumentos possíveis no discurso prático. Inclusive argumentos jurídicos imaginários. Isso vale para todos os casos de emprego de normas de inferência lógica nas argumentações jurídicas” (ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 242).

¹³⁶⁹ ASHLEY, Kevin. Evaluating the Uses of Values in a Model of Legal Reasoning, p. 32.

¹³⁷⁰ Ibid., p. 33.

¹³⁷¹ “In addition, the formalism models, and we plan to implement, argument schemes for drawing analogical inferences from cases with values. The analogies are draw in terms of intermediate legal concepts. And, the knowledge representation will be sufficiently detailed to enable the program to generate arguments about whether the analogy holds” (Ibid., p. 32).

¹³⁷² ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 243.

¹³⁷³ “. (1) Para todo x : se x é um contrato de compra e venda (F) ou om contrato semelhante ao de compra e venda (F sim), então deve ser aplicáveis a x os parágrafos 433 e seguintes, do BGB” (Id.).

.(2) (x) (Hx $\rightarrow$ F sim x)¹³⁷⁴

.(3) (x) (Hx $\rightarrow$ OGx) (1),(2)¹³⁷⁵

Por sua vez, o c) **argumentum ad absurdum** se baseia na inferência de que uma consequência jurídica (**Z**) é repelida pelo sistema normativo (é tida como ilegal ou ilegítima), de maneira que, caso a proposição normativa aplicada (**R'**) leve à consequência juridicamente ilegítima, a referida proposição será proibida (**$\neg R'$**). Observa-se que o *argumentum ad absurdum* pode ser classificado como uma espécie de “*argument from a counterexample*”¹³⁷⁶ (na classificação delineada por Kevin Ashley), pois um contra-exemplo para um argumento é uma situação que mostra que o argumento pode ter premissas verdadeiras e uma conclusão falsa. Desse modo, o *argumentum ad absurdum* pode ser formalizado através da lógica proposicional, conforme se observa da regra argumentativa (**J.17**)¹³⁷⁷:

(J.17) .(1) O $\neg$ Z

.(2) R' $\rightarrow$ Z

.(3) $\neg R'$

No mais, a regra argumentativa (**J.18**) define que as “formas de argumentos jurídicos especiais devem ser saturadas”¹³⁷⁸, o que denota o respeito à pretensão de correção.

Portanto, tendo em vista que o “**resultado da análise de um objeto complexo pode somente ser um modelo complexo**”¹³⁷⁹, percebe-se que a justificação externa exerce um papel organizativo no discurso jurídico¹³⁸⁰, o qual é necessário para o

¹³⁷⁴ “.(2) Para todo x: se x é um contrato que tem por objeto a transmissão onerosa de um local de negócios (H), então x é um contrato semelhante ao de compra e venda (F sim)” (Id.).

¹³⁷⁵ “.(3) Para todo x: se x é um contrato que tem por objeto a transmissão onerosa de um local de negócios (H), então se lhe deve aplicar os parágrafo 433 e seguintes, do BGB (1), (2)” (Id.).

¹³⁷⁶ ASHLEY, Kevin. Evaluating the Uses of Values in a Model of Legal Reasoning, p. 33.

¹³⁷⁷ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 245.

¹³⁷⁸ Id.

¹³⁷⁹ ALEXY, Robert. **Direito, Razão, Discurso**, p. 89.

¹³⁸⁰ “A argumentação prática geral pode ser necessária (1) na fundamentação das premissas normativas requeridas para a saturação das diferentes formas de argumentos, (2) na fundamentação da eleição de diferentes resultados, (3) na fundamentação e composição de enunciados dogmáticos, (4) na fundamentação dos *distinguishing* e *overruling* e (5) diretamente na fundamentação dos enunciados a serem utilizados na justificação interna” (ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 245-246).

estabelecimento das bases mínimas necessárias para a utilização da teoria da argumentação jurídica como um paradigma possível para o desenvolvimento de sistemas de Inteligência Artificial judicial.

4.4 IA e Razoabilidade: do Contexto Fático ao Teste Lógico nas Proposições Normativas Auxiliadas por Inteligência Artificial

4.4.1 O Razoável: Os Limites Fáticos e a Expressão da Razoabilidade através da IA

No presente tópico, desde já se reconhece que o tema da razoabilidade, de origem geográfica diversa e em contexto jurídico diferenciado em relação à teoria de Robert Alexy (marco teórico principal deste estudo), não se confunde com o tema da proporcionalidade, pois ambos os temas possuem elementos que são peculiares a cada um. Contudo, conforme será descrito ao longo deste subcapítulo, a razoabilidade encontra normatividade constitucional e infraconstitucional no Direito brasileiro, o que denota que a mesma não pode ser descartada como uma técnica argumentativa com o potencial de aplicabilidade na interação entre decisão judicial e Inteligência Artificial, até porque a razoabilidade e a proporcionalidade têm pressupostos de utilização diversos.

Alguns juristas podem dizer que esta obra se afastaria do pensamento de Alexy nesse ponto, o que, por outro lado, apenas demonstra que a tese se mantém fiel ao seu objetivo de investigar o entrelaçamento entre decisão judicial e IA, ainda que a compatibilização entre teorias diferentes seja um desafio num ordenamento jurídico como o brasileiro, marcado por variadas influências epistêmicas. No fundo, o debate sobre razoabilidade e Inteligência Artificial encontra eco no subcapítulo 3.3 desta obra, pois, para além da previsão constitucional e legal, existe um contexto de pluralidade epistêmica do julgador, que demanda a aplicação da razoabilidade. Após esse aviso necessário, passa-se ao tema deste subcapítulo.

No atual estágio do conhecimento produzido pela humanidade, é possível registrar uma série de importantes desenvolvimentos em diferentes campos do saber. Porém, não são todas as pessoas que precisam ou querem compreender de maneira pormenorizada e, com o respectivo domínio da linguagem técnica particular, como funciona a física quântica (ou mesmo a newtoniana), a teoria da evolução ou uma

pesquisa antropológica. Isso indica que um sistema de percepção fática não precisa inferir todo o mundo da vida à sua volta, simplesmente porque isso não é sempre computável pela mente humana em sua integralidade.¹³⁸¹

Na aplicação da razoabilidade, em suas diversas facetas, parte-se de uma interpretação da realidade, da construção interativa entre o sistema normativo jurídico e o mundo fático (ambos permeados pelo posicionamento humano decorrente do mundo axiológico). Nas abordagens da razoabilidade, as decisões são testadas e aferidas pela *régua* do mundo da vida, de modo que aquelas decisões que desafiem a lógica do que regularmente acontece no mundo fático, devem ser afastadas e tidas como irracionais, ou simplesmente irrazoáveis.

Há um primeiro nível de interação entre o mundo da vida e a norma, interação esta que passa pela *seleção da norma aplicável* a determinado fato, o que se define com maior grau de determinabilidade e é demonstrado por autores como Neil MacCormick¹³⁸² e Robert Alexy.¹³⁸³ A fixação da norma ao caso não se trata de ato simples, haja vista envolver a decomposição de conceitos linguísticos utilizados no texto normativo¹³⁸⁴ com a finalidade de sua aplicação à situação apresentada juridicamente.

No recorte do presente tópico, o enfoque está direcionado ao nível da argumentação mais complexa, denominada por Robert Alexy como justificação externa¹³⁸⁵ e, guardadas as diferenças de modelos teóricos, chamada por MacCormick de justificação de segunda ordem.¹³⁸⁶

Alexy pouco se refere aos aspectos da razoabilidade em sua Teoria da Argumentação Jurídica (sendo que as poucas referências que podem ser inferidas são implícitas), lembrando que a referida obra foi publicada originalmente em 1978, ou seja, antes de o autor expor seu pensamento sobre o método da proporcionalidade que ficou conhecido principalmente através da obra Teoria dos Direitos

¹³⁸¹ KAHNEMAN, Daniel. Op.cit., posição 613; ROVELLI, Carlo. **A Ordem do Tempo**, p. 23; DENNETT, Daniel C. **From Bacteria to Bach and Back**. London: Penguin Books, 2017, p. 336. E-book Kindle; GUSTIN, Miracy Barbosa de Sousa; DIAS, Maria Tereza Fonseca; NICÁCIO, Camila Silva. **(Re)pensando a Pesquisa Jurídica: Teoria e Prática**. São Paulo: Almedina, 2020, p. 60. E-book Kindle.

¹³⁸² MACCORMICK, Neil. **Argumentação Jurídica e Teoria do Direito**, p. 31.

¹³⁸³ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 163-164.

¹³⁸⁴ MACCORMICK, Neil. **Argumentação Jurídica e Teoria do Direito**, p. 83-92.

¹³⁸⁵ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 203-204.

¹³⁸⁶ MACCORMICK, Neil. **Argumentação Jurídica e Teoria do Direito**, p. 128-130.

Fundamentais.¹³⁸⁷ Destaca-se que, para além da origem distinta em relação à tradição de pensamento jurídico¹³⁸⁸, existem diferenças tanto das técnicas quanto do âmbito de aplicação entre o teste da razoabilidade e o procedimento da proporcionalidade (*Verhältnismässigkeit*) desenvolvido por Alexy com fundamento nas decisões do Tribunal Constitucional Federal alemão.¹³⁸⁹

Todavia, deve-se ter em conta que Alexy trata especificamente da tradição jurídica do Tribunal Constitucional alemão, ou seja, constrói sua teoria a partir de conceitos previamente desenvolvidos pela referida corte constitucional, os quais não podem ser transplantados sem mediações e traduções para a tradição jurídica brasileira. Necessário se faz evidenciar que essa última sofreu influências difusas, advindas de diversos outros sistemas jurídicos e possui desenvolvimento próprio. A partir deste posicionamento, fica demarcado que o presente tópico não ignora a especificidade do Direito brasileiro, que deve ser percebido como fenómeno específico, produto da realidade do país.

Feitas as necessárias considerações preliminares, Alexy, a partir da lista de resultados do processo de argumentação, ao comentar os elementos do discurso prático geral, indica em sua regra **(5.3)**¹³⁹⁰ que devem “ser respeitados os limites de realizabilidade faticamente dados”¹³⁹¹ e ressalta que “(5.3) exige tanto que seja

¹³⁸⁷ ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 93-103.

¹³⁸⁸ AFONSO DA SILVA, Virgílio. O proporcional e o razoável. **Revista dos Tribunais**, n. 798, 2002, p. 23-50; BARROSO, Luís Roberto. Os Princípios da Razoabilidade e da Proporcionalidade no Direito Constitucional. **BDA – Boletim de Direito Administrativo**, março/1997, p. 156-166; MATHEWS, Jud; SWEET, Alec Stone. All Things in Proportion? American Rights Doctrine and the Problem of Balancing. **Emory Law Journal**, v. 60, 2011, p. 798-875; LORENZETTO, Bruno Meneses; CLÈVE, Clèmerson Merlin. Constituição Federal, Controle Jurisdicional e Níveis de Escrutínio. **Direitos Fundamentais & Justiça**, v. 32, jul./set. 2015, p. 97-123.

¹³⁸⁹ “A broad version of reasonableness or ‘ragionevolezza’ includes in this concept everything, in principle, that can count as a criterion of practical correctness. If this comprises proportionality, and if the relation of proportionality to the further elements is clearly determined, the reasonableness test is no real alternative to proportionality analysis” (ALEXY, Robert. The Absolute and the Relative Dimension of Constitutional Rights. **Oxford Journal of Legal Studies**, 2016, p. 1-17. DOI: 10.1093/ojls/gqw013).

¹³⁹⁰ “Uma última regra deste grupo resulta o fato de que o discurso prático se desenvolve com a finalidade de resolver questões práticas realmente existentes. Certamente se pode fazer isso também por prazer, mas esta possibilidade é dependente da primeira. Isso faz com que o discurso prático cheque a resultados que possam também ser levados à prática:

(5.3) Devem ser respeitados os limites de realizabilidade faticamente dados.

A aplicação de (5.3) pressupõe um considerável conhecimento empírico.

(5.1)-(5.3) determinam diretamente o conteúdo das proposições e regras a fundamentar. Por isso as chamaremos de ‘regras de fundamentação’” (ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 181).

¹³⁹¹ Ibid., p. 257.

inteiramente possível a realizabilidade de uma norma, como também que esteja no campo do faticamente possível”¹³⁹².

Vale destacar que o recurso do intérprete à regra do discurso atinente à realizabilidade fática **(5.3)** é sempre possível, mesmo quando se trata do discurso jurídico propriamente dito¹³⁹³, pois, conforme a regra **(6.1)**, para “qualquer falante e em qualquer momento é possível passar a um discurso de teoria do discurso”.¹³⁹⁴ Para essa teoria, a regra do discurso jurídico **(J.10)** estabelece que “Todo enunciado dogmático, se posto em dúvida, deve ser fundamentado mediante o emprego, pelo menos, de um argumento prático de tipo geral”¹³⁹⁵.

Além disso, indica o necessário suporte fático para a realização de restrições justificadas em direitos fundamentais.¹³⁹⁶ Em suma, a teoria de Alexy admite a influência da facticidade na conformação da aplicação do Direito, como elemento de sua teoria discursiva baseada em razão prática e seu pensamento acerca do alcance *prima facie*¹³⁹⁷ e das restrições aos direitos fundamentais. Portanto, o método da razoabilidade, pautado na tensão entre a interpretação da realidade fática e da normatividade que respaldará a decisão jurídica, muito embora não seja expressamente adotado por Alexy em sua teoria, não é de todo incompatível com ela.

Dessa forma, para além do isolacionismo epistêmico a partir da abordagem de autor específico, o ordenamento jurídico brasileiro elegeu a razoabilidade¹³⁹⁸, tida como norma implícita no conceito constitucional de devido processo legal substantivo (art. 5º, LIV, da Constituição Federal)¹³⁹⁹, como um método de aplicação do Direito

¹³⁹² Ibid., p. 181.

¹³⁹³ Lembrando a tese de Alexy de que o discurso jurídico é um caso especial de discurso prático (Ibid., p. 187-189).

¹³⁹⁴ Ibid., p. 182.

¹³⁹⁵ Ibid., p. 231.

¹³⁹⁶ ALEXY, Robert. **Teoria dos Direitos Fundamentais**, p. 301-306.

¹³⁹⁷ Direitos de não-embaraço, direitos de não-afetação e direitos de não-eliminação.

¹³⁹⁸ “É preciso que o juiz reflita e decida expressamente sobre o problema preliminar de sua legitimação, examinando, inclusive, as possíveis consequências negativas e positivas de sua intervenção na matéria, em lugar do legislador ou do administrador” (SUNDFELD, Carlos Ari. **Direito Administrativo para Céticos**. São Paulo: Malheiros, 2012, p. 75).

¹³⁹⁹ “Diante disto, abrem-se duas linhas de construção constitucional, uma e outra conducentes ao mesmo resultado: o princípio da razoabilidade integra o direito constitucional brasileiro, devendo o teste de razoabilidade ser aplicado pelo intérprete da Constituição em qualquer caso submetido ao seu conhecimento. A primeira linha, mais inspirada na doutrina alemã, vislumbrará o princípio da razoabilidade como inerente ao Estado de direito, integrando de modo implícito o sistema, como um princípio constitucional não escrito. De outra parte, os que optarem pela influência norte-americana, pretenderão extraí-lo da cláusula do devido processo legal, sustentando que a razoabilidade das leis se torna exigível por força do caráter substantivo que se deve dar à cláusula” (BARROSO, Luís Roberto. Os Princípios da Razoabilidade e da Proporcionalidade no Direito Constitucional, p. 165).

com previsão expressa, conforme se observa no art. 8º do Código de Processo Civil, no art. 27, *caput*, da Constituição do Estado do Paraná e no art. 2º, *caput*, da Lei 9.784/1999 (Lei de Processo Administrativo), bem como no art. 20, *caput*, do Decreto-Lei 4.657/1942 (Lei de Introdução às Normas do Direito Brasileiro – alterada pela Lei 13.655/2018), o qual determina que, nos âmbitos administrativo, controlador e judicial, “não se decidirá com base em valores jurídicos abstratos sem que sejam consideradas as consequências práticas da decisão”.

No aspecto jurisprudencial, acerca do tema, a pesquisa no sistema de jurisprudência do Supremo Tribunal Federal¹⁴⁰⁰ retornou 19.249 decisões (1.324 acórdãos e 17.925 decisões monocráticas) em que o termo “razoabilidade” foi utilizado na referida Corte constitucional.

Já a Plataforma de Referências de precedentes do STF, extenso banco de dados mantido pela Fundação Getúlio Vargas (FGV)¹⁴⁰¹, indicou que, na verdade, 23.295 decisões do STF utilizaram o termo “razoabilidade” na fundamentação, sendo que as 500 decisões mais recentemente citadas¹⁴⁰² na base de dados se referem às seguintes matérias: 1) Direito Administrativo – 187; 2) Direito Processual Penal – 56; 3) Direito Processual Civil – 36; 4) Direito Penal – 25; 5) Direito Civil – 11; 6) Direito Eleitoral – 9; 7) Direito Previdenciário – 8; 8) “Outros” – 8; 9) Direito do Consumidor – 4; 10) Direito do Trabalho – 2; 11) Poder Judiciário – 1; 12) Direito da Criança e do Adolescente – 1; e 13) Entes Federativos – 1.

O banco de dados da FGV também indica qual é a decisão central na rede de referências entre os diversos julgados do Supremo Tribunal Federal sobre razoabilidade, destacando-se a Ação Direta de Inconstitucionalidade nº 2, cuja menção da razoabilidade ocorreu apenas no voto-vista do Ministro Sepúlveda Pertence:

“Estou consciente, por sua vez, de que a opção, que o problema impõe, não é apenas técnica. É de alta política constitucional, admitida **a razoabilidade jurídica das várias soluções aventadas**, a eleição, entre elas, por um Tribunal como este há de guiar-se no sentido que se lhe apresente como a

¹⁴⁰⁰ BRASIL. Supremo Tribunal Federal. **Pesquisa de Jurisprudência**. [2021]. Disponível em: <https://jurisprudencia.stf.jus.br>. Acesso em: 31 jul. 2021.

¹⁴⁰¹ FUNDAÇÃO GETÚLIO VARGAS. **Supremo em Números: Plataforma de Referências**. [2021]. Disponível em: http://precedentes.cloud.fgv.br/search?termo_busca=razoabilidade. Acesso em: 31 jul. 2021.

¹⁴⁰² A Plataforma de Referências não indica o período de corte utilizado na pesquisa.

mais adequada a efetividade da Constituição, que é o cometimento fundamental desta Casa” (destaque nosso).¹⁴⁰³

Sem rumar para o realismo jurídico, que não é o caminho por onde transita o presente tópico, a pesquisa de julgamentos do STF é campo frutífero para se analisar quais são as tendências decisórias e como se dá o processo de construção dos precedentes, o que passa pela jurimetria¹⁴⁰⁴ do uso da razoabilidade como critério de interpretação, demonstrando que a razoabilidade é elemento recorrente na construção jurisprudencial do Judiciário nacional e, por isso, tema relevante para a compreensão do Direito brasileiro.

Também, a partir do estudo de precedentes do Supremo Tribunal Federal, Humberto Ávila indica que a razoabilidade é aceita no sistema jurídico brasileiro na condição de um método de decisão, o qual possui três facetas que lhe são inerentes: razoabilidade como equidade, razoabilidade como congruência e razoabilidade como equivalência.¹⁴⁰⁵

Nessa linha, pode-se classificar as diversas concepções de razoabilidade nos seguintes testes de racionalidade para a aplicação do direito: a) teste de idoneidade

¹⁴⁰³ BRASIL. Supremo Tribunal Federal (Plenário). **Voto-vista do Ministro Sepúlveda Pertence, na Ação Direta de Inconstitucionalidade nº 2.** Possibilidade de conhecimento de Ação Direta de Inconstitucionalidade com base da teoria da inconstitucionalidade superveniente. Brasília, DF: Supremo Tribunal Federal, [2020]. Autores: Federação Nacional dos Estabelecimentos de Ensino e Presidente da República. Relator: Ministro Paulo Brossard, 06 de fevereiro de 1992. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=AC&docID=266151>. Acesso em: 10 ago. 2021.

¹⁴⁰⁴ “A Jurimetria busca observar os comportamentos humanos em função de uma ordem jurídica. A observação, em sentido lato, caracteriza a pesquisa como empírica, porque o seu objetivo é conhecer uma parcela da realidade em que vivemos. No entanto, a observação tem também um sentido estrito. Um objeto pode ser empiricamente investigado através de dois tipos de estudo: observacional ou experimental” (NUNES, Marcelo Guedes. **Jurimetria: como a estatística pode reinventar o direito.** 2ª ed. São Paulo: Thomson Reuters Brasil, 2018, p. 118).

¹⁴⁰⁵ ÁVILA, Humberto. Op. cit., p. 139-146.

lógica¹⁴⁰⁶; b) teste de excesso¹⁴⁰⁷; c) teste finalidade-norma¹⁴⁰⁸; d) teste de eliminação de hipóteses¹⁴⁰⁹; e) teste de discriminação arbitrária¹⁴¹⁰; e f) teste de conveniência/inconveniência de interesse público.¹⁴¹¹

Na prática jurídica brasileira, a tradição da razoabilidade, advinda do direito anglo-saxão, não pode simplesmente ser transplantada¹⁴¹² para o sistema jurídico brasileiro, pois tal prática poderia levar à deformação de características a ele inerentes¹⁴¹³, ainda que notavelmente influenciadas pelo sistema da *Common Law*, a partir do Código de Processo Civil de 2015.¹⁴¹⁴

Dentre os seis testes de razoabilidade referidos, tendo em conta referida realidade, pode-se afirmar que o teste de conveniência/inconveniência de interesse público citado por Neil MacCormick¹⁴¹⁵ tende a apresentar grande discricionariedade

¹⁴⁰⁶ INGLATERRA. England and Wales Court of Appeal. Associated Provincial Picture Houses Ltd. vs. Wednesbury Corporation, 1948, 1 KB 223, p. 228-230.

¹⁴⁰⁷ “O entendimento desta Corte é no sentido de que o valor estabelecido pelas instâncias ordinárias a título de indenização por danos morais pode ser revisto tão somente nas hipóteses em que a condenação se revelar irrisória ou exorbitante, distanciando-se dos padrões de razoabilidade, que não é o caso dos autos” (BRASIL. Superior Tribunal de Justiça (4. Turma). **Agravo Interno no Agravo em Recurso Especial nº 1.718.713/GO**. Acórdão de decisão que negou provimento ao pedido de revisão de valor de indenização por dano moral. Brasília, DF: Superior Tribunal de Justiça, [2020]. Recorrente: Dangelacia Teofila Constantino. Recorridos: Bradesco Auto e outros. Relator: Ministro Marco Buzzi, 26 de outubro de 2020. Disponível em: https://scon.stj.jus.br/SCON/GetInteiroTeorDoAcordao?num_registro=202001502389&dt_publicacao=12/11/2020. Acesso em: 10 ago. 2021).

¹⁴⁰⁸ “Uma regra não é aplicável somente porque as condições previstas em sua hipótese são satisfeitas. Uma regra é aplicável a um caso se, e somente se, suas condições são satisfeitas e sua aplicação não é excluída pela razão motivadora da própria regra ou pela existência de um princípio que institua uma razão contrária” (ÁVILA, Humberto. Op. cit., p. 142).

¹⁴⁰⁹ “O caráter de nossos vários tipos de caso ‘problema’ é precisamente determinado pelo fato de eles envolverem possibilidades rivais – fazer valer ou não fazer valer uma determinada deliberação, interpretar um determinado dispositivo desse modo ou daquele, tratar fatos F1, F2, F3 como manifestações ou de ‘p’” (MACCORMICK. **Argumentação Jurídica e Teoria do Direito**, p. 129).

¹⁴¹⁰ “Naturalmente, um precedente ostensivamente contraditório pode ser ‘explicado’ e ‘discriminado’ para evitar uma contradição semelhante; ou uma lei ostensivamente conflitante pode ser interpretada de uma forma que evite essa contradição. Contudo, se essas estratégias para a conciliação falharem, a exigência da coesão demandaria a rejeição de uma deliberação que de outro modo seria interessante, rejeição esta com base em seu conflito insolúvel com (a contradição de) normas válidas e estabelecidas” (Ibid., p. 135)

¹⁴¹¹ Ibid., p. 147.

¹⁴¹² “Quando o jurista se situa em face de um Direito como o brasileiro, que diz sim ao Código Civil, e a um Direito como o inglês, que diz não ao Código Civil, ele deve perceber com clareza que esses dois sistemas jurídicos são ‘incomensuráveis’ um em relação ao outro” (LEGRAND, Pierre. **Como Ler o Direito Estrangeiro**. Tradução: Daniel Wunder Hachem. São Paulo: Contracorrente, 2018, posição 223. E-book Kindle).

¹⁴¹³ “O juiz de um sistema romano-germânico como o brasileiro não pode buscar a solução dos litígios apenas nos fatos, pois deve demonstrar a incidência da legislação sobre o caso” (PUGLIESI, William Soares. **Princípios da Jurisprudência**. Belo Horizonte: Arraes Editores, 2017, p. 37).

¹⁴¹⁴ Ibid., p. 54.

¹⁴¹⁵ “Aí temos um reconhecimento explícito da possibilidade de uma colisão entre as exigências da justiça e da conveniência, com esta última, nessa ocasião, recebendo prioridade.

ao julgador, o que torna difícil sua compatibilização com o sistema brasileiro voltado à construção das decisões judiciais (o objeto do presente tópico), todavia que se apresenta valioso para o estudo das decisões administrativas pautadas em discricionariedade.

Já o teste de discriminação arbitrária pode ser relacionado à ponderação inerente ao direito à isonomia, o qual se manifesta em diversas normas constitucionais (art. 5º, *caput*, I, art. 3º, IV, art. 7º, XXXIV, art. 14, *caput*, art. 37, *caput*, II e XXI, e art. 150, II, da Constituição Federal), e é tema já bastante estudado pela doutrina brasileira¹⁴¹⁶ e de relativa normalidade de aplicação jurisprudencial.

O teste de eliminação de hipóteses consiste na seleção, dentro das diversas proposições possíveis, para a definição de um conceito normativo, qual seria o mais adequado e afinado à realidade de um determinado ordenamento jurídico ou de um determinado caso. Aqui se trata, também, de um tema clássico da interpretação jurídica, o qual fica marcado a partir da viragem linguística que tem como expoente Ludwig Wittgenstein¹⁴¹⁷, filósofo que influenciou inúmeros autores, como, por exemplo, Herbert Hart¹⁴¹⁸, o qual, por sua vez, influenciou diretamente a construção das posições de Alexy¹⁴¹⁹ e MacCormick¹⁴²⁰.

Em especial, o pensamento de MacCormick dá destaque à função classificadora que caracteriza a análise das consequências de cada hipótese de

O sentido de ressaltar a possibilidade dessa colisão está em que ela revela por que poderia ser enganoso considerar simplesmente utilitaristas os argumentos consequencialistas no direito. A argumentação consequencialista de fato trata de estabelecer que uma deliberação preferida é a melhor em vista de todas as considerações. Contudo, essa conclusão quanto à 'melhor' não é determinada por referência a uma única escala de avaliação (por exemplo, a escala do prazer - dor como no utilitarismo hedonista de um Bentham). Ela é uma sentença definitiva expressa como soma dos resultados cumulativos ou concorrentes de avaliação por meio de referência a uma série de critérios de valor, aí incluídos a 'justiça' e 'senso comum' bem como o 'proveito público' e a 'conveniência'. Algumas versões do 'utilitarismo ideal', ao contrário do 'utilitarismo hedonista', permitem critérios mais complexos em vez de mais simples no que diz respeito ao que é 'melhor' e com base nisso seria possível considerar que a justificação consequencialista no direito envolva uma forma de utilitarismo ideal" (MACCORMICK, Neil. **Argumentação Jurídica e Teoria do Direito**, p. 147).

¹⁴¹⁶ ÁVILA, Humberto. Op. cit., p. 137-138.

¹⁴¹⁷ "A expressão pressupõe as formas de todas as proposições em que pode aparecer. É a marca característica comum de uma classe de proposições" (WITTGENSTEIN, Ludwig. **Tractatus Logico-Philosophicus**, p. 147).

¹⁴¹⁸ HART, Herbert. **O Conceito de Direito**, p. 164.

¹⁴¹⁹ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 195-203.

¹⁴²⁰ "La duda razonable no es lo mismo que la toma de decisiones razonable ni tampoco es lo mismo que el cuidado razonable en la conducción. Pero hay un hilo común que une la apelación a lo 'razonable' en estos y otros casos de su uso" (MACCORMICK, Neil. **Retórica y Estado de Derecho**, posição 6112).

proposição normativa¹⁴²¹, o que leva à pergunta de qual é a diferença que determinados resultados possuem a partir da adoção de determinada proposição normativa dentre outra(s) possível(is) numa situação concreta. Assim, para MacCormick a razoabilidade dos resultados possíveis de uma interpretação¹⁴²², mediante um utilitarismo de norma¹⁴²³, apresenta-se como um filtro atinente ao desafio semântico de se trabalhar a partir de normas com abertura conceitual.¹⁴²⁴

Nessa linha, o processo de justificação da decisão ocorre mediante a atrição entre as hipóteses de aplicação da norma¹⁴²⁵ e os limites do que acontece de maneira empírica no mundo fático¹⁴²⁶, o que é inerente à razoabilidade, uma vez que em relação à lógica de justificação, “as decisões jurídicas tratam do ‘mundo real’, da mesma forma que as hipóteses científicas; e o fazem não num vazio, mas no contexto

¹⁴²¹ “Isso por sua vez significa que escolher entre possíveis deliberações opostas num dado caso envolve escolher entre o que deverá ser concebido como modelos opostos, como padrões opostos de conduta nessa sociedade. Ou bem os fabricantes de bens de consumo deveriam tomar cuidados razoáveis no preparo e acondicionamento de seus produtos, e deveriam estar sujeitos a um pedido de indenização por parte de qualquer pessoa prejudicada por negligência do fabricante a esse respeito; ou então não lhes é exigido nos termos da lei que tomem esse tipo de cuidado e nos termos da lei não são responsabilizados por falhas nesse sentido. Para considerar a sério que essa alternativa representa uma escolha verdadeira numa sociedade real, é preciso então perguntar qual é a diferença. E a resposta é que a diferença é determinada pelas diferenças que resultariam da adoção e aplicação de uma ou da outra dessas deliberações opostas numa situação social concreta” (MACCORMICK, Neil. **Argumentação Jurídica e Teoria do Direito**, p. 132).

¹⁴²² (...) “é preciso ressaltar que, como o foco dessas justificações é voltado para as consequências de deliberações contrárias (na forma de proposições universais o direito), não nas consequências para as partes particulares da decisão em particular, o utilitarismo envolvido é o ‘utilitarismo de norma’, não o ‘utilitarismo de ato’” (Ibid., p. 147-148).

¹⁴²³ “Em seu brilhante estudo pioneiro, R. A. Wasserstrom defendeu o que chamou de ‘Procedimento de Justificação em Dois Níveis’ como um bom procedimento par justificar decisões jurídicas. Com isso ele se referiu a um procedimento por meio do qual decisões particulares fossem justificadas por serem dedutíveis de normas adotadas por juízes por apresentarem melhores consequências que qualquer norma alternativa possível que também pudesse abranger a decisão” (Ibid., p. 148).

¹⁴²⁴ “*Es un dicho común el de que hay muchas cuestiones sobre las que las personas razonables pueden discrepar razonablemente. Algunas de ellas son simples diferencias de gusto personal —el baloncesto para una persona es un juego más emocionante que el críquet, pero otra prefiere el críquet por la larga y lenta creación de expectación y tensión en contraste con la acción más explosiva del baloncesto—. De gustibus non est disputandum, es absurdo tratar las diferencias de gusto como ocasiones para la disputa*” (MACCORMICK, Neil. **Retórica y Estado de Derecho**, posição 6121).

¹⁴²⁵ “O processo de experimentação é um processo de verificação de duas hipóteses contrárias como explicações. Pois uma experiência pertinente é aquela que consegue provar a falsidade de uma ou outra das previsões acerca de seu resultado que o cientista deriva das duas hipóteses contrárias. Nenhuma teoria pode jamais ser provada verdadeira por um procedimento desses. Contudo, se uma teoria for corroborada enquanto uma contrária é refutada por essa experimentação, será justificado que acatemos a primeira em lugar da segunda” (MACCORMICK, Neil. **Argumentação Jurídica e Teoria do Direito**, p. 129).

¹⁴²⁶ “Logo, um elemento na verificação diz respeito ao que acontece no mundo: qual é a prova empírica? Existe outro porém, já que a interpretação das provas envolve necessariamente o uso de pressupostos que pertencem eles mesmos à teoria científica”. (...) “Para adotar uma postura técnica, a verificação sempre implica a dependência de ‘hipóteses auxiliares’ que são elas mesmas, num dado experimento, tidas como líquidas e certas. Embora elas também possam ser objeto de experimentação direta - caso no qual outras hipóteses auxiliares estarão envolvidas” (Ibid., p. 130).

de toda uma estrutura de ‘conhecimento’¹⁴²⁷, consistente no sistema normativo do Direito.

A ocorrência de atos irrazoáveis (administrativos, legislativos ou judiciais) que podem ser objeto dos testes de razoabilidade se pauta na experiência daquilo que já é conhecido, havendo uma função heurística na argumentação por meio de elementos, *a priori*, conceituais¹⁴²⁸, os quais nem sempre serão objeto de disputa.

4.4.2 Razoabilidade e IA: Do Teste *Wednesbury* ao Teste de Idoneidade Lógica

A razoabilidade tem uma destacada função heurística¹⁴²⁹ necessária para casos subsequentes, semelhantes ou alcançados pela racionalidade decisória mediante analogia, já que as decisões baseadas em razoabilidade têm o papel de estabelecer coesão ao sistema jurídico ao eliminar interpretações que não seriam suportadas. Nesse sentido, MacCormick adota em seu pensamento concepções de a) razoabilidade consequencialista através da analogia¹⁴³⁰; b) razoabilidade mediante a aceitabilidade das consequências¹⁴³¹; e de c) desacordo razoável (valorativo) sobre a razoabilidade¹⁴³².

¹⁴²⁷ Ibid., p. 131.

¹⁴²⁸ “É preciso saber muito, antes que se possa saber sequer o que é que se está procurando – e nessa medida o conhecimento do pesquisador lhe dá ‘pistas’ em várias direções que são dignas de reflexão e experimentação. Nessa medida, mesmo as descobertas mais espantosas e de criatividade mais brilhante envolvem necessariamente uma extrapolação a partir do que já se conhece, ao longo de linhas determinadas pelo conjunto de conhecimento teórico existente” (Ibid., p. 130).

¹⁴²⁹ “Portanto, nunca verificamos hipóteses científicas simplesmente num vazio. Elas são sempre e necessariamente verificadas no contexto do todo de uma teoria com a qual sejam compatíveis e em conjunto com a qual façam sentido. E mais, com a possível exceção de descobertas revolucionárias na escala grandiosa de um Newton ou um Einstein, a realização de ‘descobertas’ – o lampejo de percepção que revela uma nova hipótese explanatória – ocorre no interior de uma estrutura de conhecimento científico” (Ibid., p. 130).

¹⁴³⁰ “(i) Esse é um modo consequencialista de argumentar, embora num sentido algo restrito. Ele considera as consequências de adotar uma deliberação por um lado ou pelo outro, pelo menos até o ponto de examinar os tipos de decisão que teriam de ser tomados em outros casos hipotéticos que poderiam ocorrer e que se enquadrariam nos termos da deliberação” (Ibid., p. 134).

¹⁴³¹ “(ii) É uma argumentação de caráter intrinsecamente avaliatório, na medida em que indaga sobre a aceitabilidade ou inaceitabilidade dessas consequências. Não existe, porém nenhuma razão para pressupor que ela envolva uma avaliação em termos de uma única escala, como, por exemplo, a escala benthamita de agregados supostamente mensuráveis de prazeres e dores. Os juízes caracteristicamente referem-se a critérios como os da ‘justiça’, ‘senso comum’, ‘interesse público’ e ‘conveniência’ ou ‘praticidade’, ao avaliar os prós e os contras de deliberações determinadas. Não se deveria pressupor sem provas que tudo isso acabe se reduzindo à mesma coisa. Por esse motivo, deveríamos ter a cautela de evitar chamá-la de argumentação ‘utilitária’, se bem que ela apresente acentuada semelhança com o que às vezes é denominado ‘utilitarismo da norma ideal’” (Id.).

¹⁴³² “(iii) Ela é, pelo menos em parte, subjetiva. Os juízes, ao avaliar as consequências de possíveis deliberações opostas, podem atribuir pesos diferentes a critérios diferentes de avaliação; podem divergir quanto ao grau de injustiça percebida, ou do inconveniente previsto que resultará da adoção

Dessa forma, levando em conta conjuntamente os três aspectos referidos, é possível deduzir “que a justificação de segunda ordem diga respeito ao ‘que faz sentido no mundo’, na medida em que ela envolve argumentos consequencialistas que são essencialmente de caráter avaliatório”¹⁴³³, ou seja, tratam-se de argumentos com carga subjetiva influenciados pelo intérprete institucionalmente inserido.

Assim, as consequências das decisões judiciais são precedidas de determinadas condições fixadas no ordenamento normativo¹⁴³⁴, as quais se entrelaçam com a realidade fática¹⁴³⁵ no processo de aplicação do Direito. Nesse sentido, a estrutura de conhecimento do mundo real, para além do mundo normativo, se expressa através de ontologias que podem ser aproveitadas pelo intérprete para a tomada de decisão.

Nessa linha, a razoabilidade decorre da ideia de universalidade¹⁴³⁶, a fim de que haja um nível de análise consequencialista inerente à busca da decisão razoável que rejeite a arbitrariedade decisória, garantindo um processo de deliberação por meio de critérios racionais/razoáveis que tenham nível de generalidade das consequências a serem testadas.¹⁴³⁷

ou rejeição de uma determinada deliberação. De modo não surpreendente, eles divergem, às vezes de modo acentuado e veemente quanto à sua opinião final, considerados todos os aspectos, sobre a aceitabilidade ou inaceitabilidade de uma deliberação em exame. É nesse ponto que chegamos aos fundamentos das preferências de valor que moldam nosso raciocínio mas que não são demonstráveis por ele. Nesse nível, pode haver simplesmente diferenças insolúveis de opinião entre pessoas racionais e de boa vontade” (Id.).

¹⁴³³ Id.

¹⁴³⁴ “As deliberações legais são normativas. Elas não relatam mas fixam padrões de comportamento. Não descobrem as consequências de determinadas condições, mas ordenam quais consequências hão de se seguir a determinadas condições. Não apresentam um modelo do mundo, mas apresentam um modelo para o mundo” (Ibid., p. 132).

¹⁴³⁵ “A argumentação empírica serve tanto diretamente para a justificação dos enunciados empíricos empregados na justificação interna, como também para a justificação de enunciados empíricos empregados na justificação externa de enunciados não-empíricos” (ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 204).

¹⁴³⁶ “Muito bem que se diga que o princípio da justiça formal requer que o juiz, expressa ou implicitamente, formule e teste alguma deliberação possível para o caso e para casos semelhantes a fim de justificar uma decisão no caso. No entanto, (como a distinção traçada anteriormente entre ‘universalidade’ e ‘generalidade’ indica), uma deliberação adequadamente ‘universal’ poderia ser atingida em qualquer dos ‘níveis de generalidade’ de Stone” (MACCORMICK, Neil. **Argumentação Jurídica e Teoria do Direito**, p. 151).

¹⁴³⁷ (...) “mesmo que a argumentação consequencialista possa demonstrar fundamentos racionais para a escolha entre deliberações possíveis no caso, existe algum fundamento racional ou mesmo razoável para escolher qual deliberação e em que nível de generalidade será testada por aquele procedimento? Se não existir, mais uma vez devemos concluir que um processo ostensivamente racional é no fundo nada mais do que arbitrário. Enfrentamos a questão: que limites, se é que existem, podem reger a escolha de deliberações a serem testadas; e como, em qualquer caso, podem os juízes começar a estruturar alguma deliberação adequada para se encaixar non caso concreto quando está aberto um leque tão amplo de possibilidades” (Ibid., p. 152).

Dessa forma, para a delimitação do presente tópico e tendo em vista os aspectos que se propõem ao aprofundamento, a partir desse ponto o enfoque ocorrerá em relação à concepção de razoabilidade a partir do teste de racionalidade relativo à idoneidade lógica¹⁴³⁸ (destacando-se, todavia, a igual importância do teste de excesso e do teste finalidade-norma para a definição da razoabilidade no Direito brasileiro).

O teste de idoneidade lógica tem origem na ideia de razoabilidade advinda do direito inglês, no chamado teste de *Wednesbury*¹⁴³⁹, o qual se constitui num teste de irrazoabilidade¹⁴⁴⁰, como no “exemplo da professora ruiva, dispensada porque tinha cabelos ruivos” (tradução livre)¹⁴⁴¹ lembrado por Lorde Greene no julgamento do *leading case* referido.

No teste de idoneidade lógica inerente à razoabilidade, há uma vinculação à ideia de que existem decisões tão irracionais que quase podem ser descritas como feitas mediante má-fé¹⁴⁴², ainda que esse aspecto seja objetivado ao se utilizar o conceito de uma “autoridade razoável”¹⁴⁴³ para fins de definição de um padrão de conduta racional inerente à tomada de decisões públicas, pois “se uma decisão sobre uma matéria competente for tão irrazoável que jamais ela poderia ter vindo de uma autoridade razoável, então os tribunais podem interferir”.¹⁴⁴⁴ Ao fundamentar a decisão da *Court of Appeal* Lorde Greene assim expressou o teste de irrazoabilidade:

“Não o considero de forma alguma discordando da visão que me aventurei a expressar, de que a tarefa do tribunal não é decidir o que pensa ser razoável, mas decidir se o que está *prima facie* dentro do poder da autoridade local é uma condição que nenhuma autoridade razoável, agindo nos quatro cantos de sua competência territorial, poderia ter decidido impor” (tradução nossa).¹⁴⁴⁵

¹⁴³⁸ INGLATERRA. Op. cit.

¹⁴³⁹ Id.

¹⁴⁴⁰ “E esse teste da irrazoabilidade, conhecido também como teste *Wednesbury*, implica tão somente rejeitar atos que sejam excepcionalmente irrazoáveis. Na fórmula clássica da decisão *Wednesbury*: “se uma decisão [...] é de tal forma irrazoável, que nenhuma autoridade razoável a tomaria, então pode a corte intervir”. Percebe-se, portanto, que o teste sobre a irrazoabilidade é muito menos intenso do que os testes que a regra da proporcionalidade exige, destinando-se meramente a afastar atos absurdamente irrazoáveis.” (AFONSO DA SILVA, Virgílio. Op. cit.).

¹⁴⁴¹ INGLATERRA. Op. cit.

¹⁴⁴² Id.

¹⁴⁴³ Id.

¹⁴⁴⁴ Tradução livre de: “*It is true to say that, if a decision on a competent matter is so unreasonable that no reasonable authority could ever have come to it, then the courts can interfere*” (Id.).

¹⁴⁴⁵ Tradução livre de: “*I do not read him as in any way dissenting from the view which I have ventured to express, that the task of the court is not to decide what it thinks is reasonable, but to decide whether*

Ainda ao final de sua fundamentação, Lorde Greene destacou que o poder de análise dos atos irrazoáveis proferidos por autoridades administrativas decorre da obrigação de “uma autoridade judicial que se preocupa, e apenas se preocupa, em ver se a autoridade local infringiu o direito agindo para além dos poderes que o Parlamento lhes confiou”.¹⁴⁴⁶

A origem da razoabilidade no direito inglês, por evidente, ocorreu no direito administrativo, embora, com o desenvolvimento jurisprudencial e doutrinário, a razoabilidade tenha alcançado o status de elemento intrínseco do modelo de *judicial review* inglês.

A partir de uma análise dos argumentos utilizados no caso *Wednesbury*, evidencia-se o tipo padrão de irrazoabilidade, de índole inerente à ausência de conexão lógica (congruência) entre o fato e a consequência jurídica aplicada, posição também indicada por MacCormick.¹⁴⁴⁷ Vê-se que MacCormick adota um conceito de autoridade razoável¹⁴⁴⁸, embora seja cético quanto à existência de pessoas realmente “razoáveis”¹⁴⁴⁹, inclusive porque tais pessoas podem ter um desacordo razoável sobre a própria razoabilidade.¹⁴⁵⁰

what is prima facie within the power of the local authority is a condition which no reasonable authority, acting within the four corners of their jurisdiction, could have decided to impose” (Id.).

¹⁴⁴⁶ Tradução livre de: “*The power of the court to interfere in each case is not as an appellate authority to override a decision of the local authority, but as a judicial authority which is concerned, and concerned only, to see whether the local authority have contravened the law by acting in excess of the powers which Parliament has confided in them*” (Id.).

¹⁴⁴⁷ “*De nuevo, lo que justifica el recurso al requisito de razonabilidad es la existencia de una pluralidad de factores que requieren ser evaluados en cuanto a su pertinencia para un foco común de interés (en este caso, una decisión que debe tomar un organismo público con fines públicos). La irrazonabilidad consiste en ignorar algún factor o algunos factores pertinentes, o en tratar como pertinente lo que debe ser ignorado*” (MACCORMICK, Neil. **Retórica y Estado de Derecho**, posição 6238).

¹⁴⁴⁸ “*La persona razonable tiene la virtud de la prudencia y hace uso de ella en la acción. Es una virtud incompatible con el fanatismo o la apatía, pero ocupa un lugar medio entre ellas, al igual que lo hace entre la excesiva cautela y la excesiva indiferencia al riesgo. Las personas razonables tienen en cuenta los riesgos previsibles, pero tomando en consideración posibilidades o probabilidades serias, no posibilidades remotas o caprichosas*” (Ibid., posição 6083).

¹⁴⁴⁹ “*Las personas perfectamente razonables serían sin duda ejemplos irreales de virtud. Existen pocas. Las personas ordinarias no lo son, pero la mayoría son razonables en algunos momentos y algunas son razonables la mayor parte del tiempo*” (Ibid., posição 6103).

¹⁴⁵⁰ “*Una evaluación de este tipo tiene un obvio aspecto subjetivo. Evaluadores diferentes se formarán opiniones diferentes sobre los beneficios de la actividad, la probabilidad de los riesgos, la seriedad de los riesgos, los costes de las medidas de seguridad y su probable eficacia. Es notorio que personas razonables que traten con imparcialidad una cuestión compleja pueden estar razonablemente en desacuerdo sobre la respuesta razonable. En ese caso puede que todo se reduzca a la cuestión de quién es el juez con mayor autoridad, pues la autoridad de esa persona prevalecerá. Eso sugiere un grado de subjetividad, ya que diferentes sujetos que consideren la misma cuestión pueden estar en desacuerdo habitualmente (no estar en desacuerdo siempre o invariablemente) sobre cuál es la respuesta correcta*” (Ibid., posição 5993).

Assim, a fórmula *Wednesbury* do teste de idoneidade lógica (“se uma decisão (...) é de tal forma irrazoável, que nenhuma autoridade razoável a tomaria, então pode a corte intervir”¹⁴⁵¹) pode ser relacionada à uma concepção de justiça compartilhada da sociedade, objetivada por meio de razões perceptíveis aos cidadãos dessa sociedade, o que passa pela consideração dos aspectos empíricos¹⁴⁵² do caso.

Nesse caminho, um desafio na aplicação da razoabilidade é a conjugação de posições diferentes do que seja razoável¹⁴⁵³. Assim, o teste de idoneidade lógica leva em consideração o fato de que a decisão tomada deve fazer sentido no mundo fático-jurídico.¹⁴⁵⁴ A partir disso, o que caracteriza a fórmula *Wednesbury* é que um julgamento qualificado como irrazoável indica um grau elevado de equívoco, permeado por um elemento de má-fé objetiva em que a decisão marcada pela irrazoabilidade tem algum tipo de “perversidade, arbitrariedade, capricho ou absurdidade que a retira da classe de um mero exercício errôneo de discricionariedade”¹⁴⁵⁵, o que indica que tem sido aceita uma “intervenção judicial que

¹⁴⁵¹ INGLATERRA. Op. cit.

¹⁴⁵² “Primeiro, a razoabilidade é utilizada como diretriz que exige a relação das normas gerais com as individualidades do caso concreto, quer mostrando sob qual perspectiva a norma deve ser aplicada, quer indicando em quais hipóteses o caso individual, em virtude de suas especificidades, deixa de se enquadrar na norma geral. Segundo, a razoabilidade é empregada como diretriz que exige a vinculação das normas jurídicas com o mundo ao qual elas fazem referência, seja reclamando a existência de um suporte empírico e adequado a qualquer ato jurídico, seja demandando uma relação congruente entre a medida adotada e o fim a que ela pretende atingir. Terceiro, a razoabilidade é utilizada como diretriz que exige a relação de equivalência entre duas grandezas” (ÁVILA, Humberto. Op. cit., p. 139).

¹⁴⁵³ “*El problema puede tener que ver con el carácter ‘procedimental’ de la deliberación razonable. A la luz de los valores, intereses y propósitos humanos, uno debe considerar todo lo que sea pertinente, y asumir una postura imparcial al atribuir un peso o una importancia relativa a los diferentes valores o intereses contextualmente pertinentes. Pero diferentes personas pueden discrepar sobre cómo exactamente atribuyen tal peso y llevan a cabo el balance. Puede que haya errores evidentes de parcialidad, o anomalías patentes en la ponderación diferencial, pero, más allá de eso, es difícil o imposible mostrar que un enfoque sea superior a otro. Siempre y cuando las personas eviten las veleidades y los caprichos, y manifiesten una constancia decente en el juicio a lo largo del tiempo, mientras permanecen abiertas a la revisión de su opinión a la luz de argumentos razonados, no dejan de ser razonables simplemente porque tengan una opinión diferente de la mía, o de la de usted*” (MACCORMICK, Neil. **Retórica y Estado de Derecho**, posição 6131-6142).

¹⁴⁵⁴ “Em termos simples, as decisões jurídicas devem fazer sentido no mundo e devem também fazer sentido no contexto do sistema jurídico. Em nossos casos-problema, elas devem ser baseadas em deliberações que façam sentido no contexto do sistema jurídico. E, exatamente como a justificação científica envolve a verificação de uma hipótese em comparação com outra e a rejeição da que não passar nos testes pertinentes, assim também (como argumentarei) a justificação de segunda ordem no direito envolve a verificação de possíveis deliberações reciprocamente opostas, por meio da comparação de uma com a outra, bem como a rejeição daquelas que não satisfizerem testes pertinentes – sendo que estes tratam do que faz sentido no mundo e do que faz sentido no contexto do sistema” (MACCORMICK, Neil. **Argumentação Jurídica e Teoria do Direito**, p. 131).

¹⁴⁵⁵ Tradução livre parcial do seguinte trecho: “*The crucial distinction, in regard to containing the Wednesbury formula, is that between a mistaken and an unreasonable exercise of judgment. Although an element of mala fides is not indispensable, the decision impeached must contain some quality of perversity, arbitrariness, caprice or absurdity which removes it from the class of a merely erroneous*

deve ser forte ou mesmo esmagadora”.¹⁴⁵⁶ Nessa linha, há a possibilidade de descrição da razoabilidade apenas em termos “grossos” (com médio grau de determinabilidade), como lembra MacCormick.¹⁴⁵⁷

Os vícios de perversidade, arbitrariedade, capricho ou absurdidade, na forma da construção historicamente delineada pela *Court of Appeal*, desde 1948, traduzem ponto de ligação entre o direito inglês e o direito brasileiro, na medida em que tais vícios são rejeitados pelo ordenamento jurídico nacional mediante uma ampla gama de outras normas aplicáveis, em reforço à razoabilidade, como os princípios da impessoalidade, da moralidade e da eficiência (art. 37, *caput*, da Constituição Federal, o art. 8º do Código de Processo Civil e art. 2º da Lei 9.784/1999), por exemplo.

Todavia, percebe-se que, com a evolução da aplicação da razoabilidade na *Court of Appeal*, ficaram caracterizados os dois estágios do teste *Wednesbury*. No primeiro estágio¹⁴⁵⁸, dos anos 1980 até meados de 1990, havia maior deferência às escolhas das autoridades públicas, reservando a intervenção judicial pelo argumento da irrazoabilidade para casos mais graves (distinção forte entre o ato irrazoável e o ato meramente equivocado¹⁴⁵⁹). No primeiro estágio do teste de *Wednesbury* os integrantes da *Court of Appeal* “mostraram grande sensibilidade à necessidade de evitar escrupulosamente até mesmo a sugestão de usurpar a substância do poder

exercise of discretion” (PEIRIS, G. L. *Wednesbury Unreasonableness: The Expanding Canvas. Cambridge Law Journal*, v. 46, n. 1, March 1987, p. 53-82).

¹⁴⁵⁶ “*The courts, therefore, have been at pains to stress that, in the absence of anything akin to appellate jurisdiction, the case for judicial intervention must be strong or even overwhelming*” (Id.).

¹⁴⁵⁷ “*El concepto de lo ‘razonable’ es usado con frecuencia y en contextos muy diversos por los abogados y los legisladores en estas tradiciones jurídicas. En el espectro que va de lo puramente descriptivo a lo puramente evaluativo, ‘razonable’ parece encontrarse más cerca del polo evaluativo que del descriptivo, aunque eso no significa que no haya ningún elemento descriptivo en él. De hecho, puede clasificarse propiamente entre los términos evaluativos que los filósofos han denominado ‘gruesos’, a diferencia de conceptos finos tales como ‘correcto’ o ‘bueno’*” (MACCORMICK, Neil. *Retórica y Estado de Derecho*, posição 5983).

¹⁴⁵⁸ “*An analysis of these cases allows the division of the English case-law into two distinct stages. In the first stage, from the 1980s until mid-1990s, the courts showed enormous deference to health authorities’ decisions for epistemic and constitutional reasons. They considered medical/scientific assessments and allocative decisions to be beyond the reach of judicial scrutiny and thus that the courts should only intervene if the administrative decision was so unreasonable that no reasonable authority could ever have come to it.*” (WANG, Daniel Wei Liang. *From Wednesbury Unreasonableness to Accountability for Reasonableness. Cambridge Law Journal*, v. 76, n. 3, 2017, p. 642-670).

¹⁴⁵⁹ “*In Hincks, Walker, Collier and Seale the courts restrained themselves to a minimal level of scrutiny of the allocative choices and trusted health authorities to make the best decisions. The court’s reasoning was straightforward: resources are scarce, not all health needs can be met, and thus rationing is necessary; and health authorities are best able to do this. On this basis, the courts determined that they had no reason to review health authorities’ decisions unless they were “Wednesbury unreasonable”, ie so absurd or outrageous in their defiance of logic or morality that no reasonable person addressing the question would have come to the same conclusion*” (Id.).

discrecionalário de seu legítimo titular”¹⁴⁶⁰, havendo maior deferência à separação dos Poderes, o que implica num controle de constitucionalidade, de atos legislativos e administrativos, que pode ser classificado como fraco.¹⁴⁶¹

Já no segundo estágio¹⁴⁶², a partir de meados da década de 1990, adotou-se um criterioso escrutínio das razões utilizadas para a formulação de atos do poder público, passando da exigência de critério de extrema injustiça para o reconhecimento de um ato como irrazoável (a absurdidade, por exemplo) para uma análise lógico-analítica da congruência entre as razões dadas para a edição do ato do poder público e seu resultado prático no mundo.

Logo, percebe-se que a segunda fase de desenvolvimento do teste de *Wednesbury* implicou na exigência da idoneidade lógica, mediante o respeito à base fática que recebe a encapsulada significação argumentativa da decisão jurídica, devendo essa respeitar aquela, o que se traduz numa sujeição das decisões jurídicas como um todo ao referido critério central do teste de razoabilidade. Em suma, não se exige que uma decisão seja absurda para ser caracterizada como irrazoável, pois, para que isso ocorra, basta que a decisão seja ilógica.

Tal espécie de análise leva a um consequencialismo advindo da necessidade de respeito ao próprio ordenamento jurídico e ao caráter institucionalista da democracia (o conceito de “autoridade razoável” do caso *Wednesbury* leva a isso), análise consequencialista esta que fortalece o Direito como a base das decisões racionais (razoáveis) que devem nortear a vida em sociedade.

O presente tópico fez uma análise desde a previsão da razoabilidade na legislação brasileira, passando pela verificação da tradicional utilização das concepções de razoabilidade na jurisprudência do Supremo Tribunal Federal.

Além disso, a partir do estudo desenvolvido, foi possível se identificar seis facetas diferentes advindas da concepção de razoabilidade, as quais se traduzem nos testes de idoneidade lógica, de excesso, de análise finalidade-norma, de eliminação

¹⁴⁶⁰ Tradução livre de: “*Judges, accordingly, have shown great sensitivity to the need for avoiding scrupulously even the suggestion of usurping the substance of discretionary power from its rightful donee*” (PEIRIS, G. L. Op. cit.).

¹⁴⁶¹ WALDRON, Jeremy. The Core of the Case Against Judicial Review. **Yale Law Journal**, v. 115, n. 6, 2006.

¹⁴⁶² “*In the second stage, from the mid-1990s onwards, the courts started judging in a way that, during the first stage, they affirmed they should not. They started scrutinizing reasons and procedures to an extent that authorities were required to act ‘synoptically’, namely to demonstrate that they did the best possible job in gathering facts, evaluating alternatives and articulating the values that had led to their decisions*” (WANG, Daniel Wei Liang. Op. cit.).

de hipóteses, de discriminação arbitrária e de conveniência/inconveniência de interesse público.

A partir da verificação da existência das diferentes vertentes da razoabilidade, mediante respeito ao recorte do presente tópico, destaca-se o teste de razoabilidade como idoneidade lógica, evidenciando-se a origem no caso *Wednesbury* (testes de perversidade, arbitrariedade, capricho e absurdidade) e sua evolução para a ideia de lógica racional na aplicação do Direito, mediante a congruência entre as medidas adotadas pelo poder público (legislativas, administrativas ou judiciais) e a base fática a ser emoldurada pela interpretação jurídica.

Sendo assim, levando-se em conta a expressa previsão da razoabilidade na legislação (art. 8º do Código de Processo Civil, no art. 27, *caput*, da Constituição do Estado do Paraná e no art. 2º, *caput*, da Lei 9.784/1999), inclusive de maneira destacada em relação à proporcionalidade, é possível a clara identificação da razoabilidade e de seu desenvolvimento teórico.

As tentativas de absorver a razoabilidade no procedimento de proporcionalidade, ou de abandonar a razoabilidade sob a alegação de que a proporcionalidade já resolve todas as situações tratadas pela primeira, não encontram respaldo, pois o teste de idoneidade lógica típico da razoabilidade (pautado numa concepção de necessidade de boa-fé objetiva da “autoridade razoável” que tem o dever de atuar de maneira correta no dado contexto fático) é acionado através de um mecanismo próprio.

Portanto, a razoabilidade se apresenta como forma de intervenção judicial que visa extirpar do direito decisões que causem tensão insustentável no plano do mundo fático. Com tal afirmação, não se ignora a primeira etapa interpretativa que é estabelecida na própria compreensão do mundo da vida, bem como os significados podem ser reconstruídos com o passar do tempo. Para isso, não há alternativa senão a busca dentro da sociedade e do momento histórico compartilhado para que se possam fixar critérios interpretativos. Cada sistema jurídico, por sua vez, acaba por funcionar como caixa de ressonância das respectivas sociedades que ordenam e reproduzem, com diferentes temporalidades, os significados compartilhados por uma comunidade política.

5 ASPECTOS DO DESENVOLVIMENTO DE INTELIGÊNCIA ARTIFICIAL JUDICIAL: A INFLUÊNCIA DA PRETENSÃO DE CORREÇÃO EM SISTEMAS DE IA JURÍDICA

Tentar resumir as diversas possibilidades de aplicação da Inteligência Artificial ao Direito extrapolaria os limites desta obra. Devido a esse motivo, antes de adentrar nos aspectos mais exploratórios das técnicas de IA, é o momento de destacar que o tema dos algoritmos de IA e aprendizado de máquina não será trabalhado em termos matemáticos, estatísticos ou mediante um estudo aprofundado de Ciência de Dados e Ciência da Computação.

A referida escolha se volta à necessidade de compreensão dos principais aspectos do aprendizado de máquina (*machine learning*) que são úteis ao estudioso do mundo do Direito, os quais são necessários para o delineamento da aplicação da tese central desta obra, no sentido de como o juiz ciborgue atua com o apoio dos sistemas de IA e qual é o potencial destes sistemas na estruturação do duplo canal de correção.

Ou seja, conforme este Capítulo, demonstra-se o contexto técnico da IA aplicada ao Direito, de maneira a relacionar isso com as categorias jurídicas trabalhadas ao longo dos demais Capítulos, especialmente para que haja a percepção de como o desenvolvimento de sistemas de Inteligência Artificial contribui para a prática de decisões judiciais que possam levar em conta o duplo canal de correção.

Logo, frente à limitação que se impõe no recorte desta obra¹⁴⁶³, a presente pesquisa não pretende esgotar o tema das técnicas de Inteligência Artificial que podem ser aplicadas ao Direito e à argumentação jurídica¹⁴⁶⁴, embora contribua com o tema numa abordagem que parte da Teoria do Direito e do Direito Constitucional.

¹⁴⁶³ Não se almeja vencer toda a temática das técnicas de desenvolvimento de inteligência artificial, o que nem mesmo obras de peso da Ciência da Computação se propõem. O que se busca é um mapeamento resumido das técnicas de desenvolvimento de IA e dos potenciais de sua relação com a decisão judicial e com a teoria da argumentação jurídica. Dessa forma, o capítulo se justifica como um esforço no sentido de melhorar a compreensão da comunidade jurídica a respeito do tema, o que, por sua vez, **contribui para a compreensão da própria tese**, uma vez que se fala da IA que é realidade, aquela que pode efetivamente auxiliar o ser humano na tomada de decisões racionais.

¹⁴⁶⁴ Para além do estudo que será desenvolvido nos subcapítulos a seguir, caso se busque um aprofundamento técnico em todas as abordagens de Inteligência Artificial aplicáveis ao Direito (inclusive para além do Judiciário), podem ser lidas, no mínimo, as seguintes obras: ASHLEY, Kevin D. **Artificial Intelligence and Legal Analytics**; BENCH-CAPON, Trevor. **Knowledge-Based Systems and Legal Applications**. London: Academic Press, 2015; SARTOR, Giovanni. Legal Concepts as Inferential Nodes and Ontological Categories. **Artificial Intelligence and Law Review**, Dordrecht, v. 17, p. 217-251, 2009; ROVER, Aires José. **Informática no Direito**; ELIOT, Lance. Op. cit.; WALTON,

Ademais, destaca-se que o debate sobre a representação do Direito em sistemas computacionais e a necessária observância dos fatores (*factors*) inerentes ao suporte fático foram tratados no subcapítulo 4.1, elementos conceituais que se adequam aos limites desta pesquisa. A partir disso, a estruturação dos modelos algorítmicos em si e a programação dos sistemas são temas que podem se aproveitar dos estudos desenvolvidos nesta obra, mas que não fazem parte do recorte epistêmico adotado para a explicação das categorias teóricas aqui desenvolvidas.

Após essa demarcação da temática, passa-se aos principais aspectos ainda não abordados no que se refere ao desenvolvimento de Inteligência Artificial aplicada ao Poder Judiciário.

5.1 Aspectos Gerais da IA Judicial: Uma Visão Para Além dos Robôs

De início, destaca-se que o presente tópico possui a função de demonstrar a importância do tema da Ciência da Computação, através da programação, da Ciência de Dados e da Inteligência Artificial, para o auxílio na interpretação do Direito através do duplo canal de correção advindo da via comunicacional humano-máquina.

Nesse caminho, é necessário, desde já, destacar que o conceito de **algoritmo** remete a uma sequência finita de passos utilizada para resolver um problema.¹⁴⁶⁵ Todavia, uma das nuances mais marcantes do Direito é que ele pode ser classificado

Douglas. **Argumentation Methods for Artificial Intelligence in Law**. Berlin: Springer, 2005; PRAKKEN, Henry. AI & Law, Logic and Argument Schemes. **Argumentation**, v. 19, p. 303-320, 2005. DOI: 10.1007/s10503-005-4418-7; HARTMANN PEIXOTO, Fabiano. **Inteligência Artificial e Direito**. Além disso, também podem ser lidos os artigos produzidos por diversos autores nos mais de trinta anos de atividade da *Artificial Intelligence and Law Review*, revista que pode ser acessada através do seguinte endereço: www.springer.com/journal/10506.

¹⁴⁶⁵ “Falar de algoritmos para nossas vidas pode parecer uma estranha justaposição. Para muita gente, a palavra ‘algoritmo’ evoca as misteriosas e inescrutáveis maquinações de grandes dados, grandes governos e grandes negócios — cada vez mais uma parte da infraestrutura do mundo moderno, mas dificilmente uma fonte de sabedoria prática ou um guia para as questões humanas. Mas um algoritmo é apenas uma sequência finita de passos que se usa para resolver um problema, e algoritmos são muito mais amplos — e muito mais antigos — do que o computador. Muito antes de serem usados por máquinas, os algoritmos eram usados por pessoas.

A palavra ‘algoritmo’ vem do nome do matemático persa *al-Khwārizmī*, autor de um livro do século IX sobre técnicas para fazer matemática à mão. (Seu livro intitulou-se *al-Jabr wa'l-Muqābala* — Livro compêndio sobre cálculo por restauração e balanceamento —, e o ‘*al-Jabr*’ do título é, por sua vez, a fonte de nossa palavra ‘álgebra’.) No entanto, os primeiros algoritmos matemáticos conhecidos precedem até mesmo a obra de *al-Khwārizmī*: um tablete de barro sumério com 4 mil anos encontrado perto de Bagdá descreve um esquema para uma longa operação de divisão” (CHRISTIAN, Brian; GRIFFITHS, Tom. Op. cit., posição 71-82).

em problemas solucionáveis através de passos bem determinados e, ao mesmo tempo, sofre da sua natural entropia linguística, política e fática.

Essa dualidade do Direito, por si só, não impede a parcial representação computacional do funcionamento de sistemas jurídicos, uma vez que a plasticidade da Inteligência Artificial jurídica é constituída por algoritmos, base de dados e modelos de IA.

A partir disso, o presente tópico conta um pouco da jornada de Direito através da Ciência da Computação e da Ciência de Dados, uma jornada que, embora não seja obrigatória, é necessária para falar da relação entre Direito e Inteligência Artificial de maneira serena e com busca pela precisão técnica, o que afasta as tentações alarmistas ou os discursos de suposto arrebatamento tecnológico.

Para o profissional do Direito, é bastante desafiador sair de seu lugar comum, do seu jargão e dos signos identificadores de sua “tribo” e buscar compreender a técnica necessária para a construção de Inteligência Artificial jurídica. A partir dessa experiência, pode ser vislumbrado que o futuro debruçar sobre os estudos de tecnologia e Inteligência Artificial tem o potencial de levar a comunidade jurídica (e parcela dos juristas) à diminuição dos excessos em relação às expectativas trazidas pelos defensores do arrebatamento tecnológico e ao abrandamento do medo paralisante que assombra os alarmistas e os negacionistas da IA jurídica.

Logo, em razão dos desafios contemporâneos, a Teoria do Direito, que já era aberta para discussões a respeito de lógica representacional voltada à interpretação jurídica e à argumentação, passa a incorporar as discussões sobre IA. Para o desenvolvimento do modelo de juiz ciborgue, esse diálogo/incremento leva a uma Teoria do Direito que é influenciada pela forma de funcionamento da Inteligência Artificial.

Nesse rumo, alguns autores trouxeram ideias sobre a aplicação de robôs no Direito, o que deve ser contextualizado no âmbito da presente obra. Um exemplo disso se verifica no modelo em que há uma atuação tímida do julgador humano na figura de revisor de “um ato completamente automatizado, através do qual se decide um processo judicial”¹⁴⁶⁶ por meio de resultados vinculativos em que se elimina “do

¹⁴⁶⁶ BOEING, Daniel Henrique Arruda; DA ROSA, Alexandre Moraes. **Ensinando um Robô a Julgar: Pragmática, Discricionariedade, Heurísticas e Vieses no Uso de Aprendizado de Máquina no Judiciário**. Florianópolis: EMais Academia, 2020, p. 101.

processo decisório completamente o componente humano”.¹⁴⁶⁷ Esse modelo se baseia nos didáticos conceitos de “Robô-Classificador”¹⁴⁶⁸, “Robô-Relator”¹⁴⁶⁹ e “Robô-Julgador”¹⁴⁷⁰ defendidos por Daniel Boeing e Alexandre Moraes da Rosa. Contudo, um modelo de grande robotização das decisões é problemático, haja vista que não leva em consideração a possibilidade de dupla via de comunicação entre o sistema de IA judicial e o julgador humano em um processo decisório dinâmico nascido do duplo canal de correção (vide o subcapítulo 1.6). Com pensamento similar, pode ser observada a defesa da robotização judicial realizada por Rômulo Soares Valentini.¹⁴⁷¹

Também pode ser referido o pensamento de Tania Sourdin, que, para além de todos os desafios que enumera, reconhece a possibilidade de implementação de sistemas de Inteligência Artificial que se limitem a auxiliar a tomada de decisão¹⁴⁷²

¹⁴⁶⁷ Ibid., p. 109.

¹⁴⁶⁸ “Robô-Classificador: Um primeiro tipo, doravante, denominado *classificador*, tem por função primordial encontrar materiais úteis para que os humanos fundamentem suas decisões. Tais materiais compreendem, por exemplo, dispositivos normativos, precedentes judiciais e modelos de documentos que servirão de base para direcionar o pronunciamento judicial” (Ibid., p. 95).

¹⁴⁶⁹ “Robô-Relator: Uma segunda forma de utilização do aprendizado de máquina, aqui apelidada de *relatora*, diz respeito a extrair e condensar informações relevantes de um ou mais documentos, o que pode ser utilizado para diferentes fins. Para tanto, ela deve ser capaz não apenas de encontrar documentos similares, mas ir mais a fundo em sua estrutura, diferenciando, em cada peça processual, aquilo que se refere à descrição de fatos, textos-legais, jurisprudências colacionadas e estruturas argumentativas” (Ibid., p. 98).

¹⁴⁷⁰ “Robô-Julgador: O terceiro e último tipo de uso do aprendizado de máquina no Judiciário, a saber, o robô-julgador, apresenta características muito próximas do modelo anterior, no que diz respeito às suas funcionalidades. Sua diferença reside principalmente no tratamento que se dá ao resultado gerado pelo algoritmo, que será tido como a própria decisão judicial. Aplica-se principalmente às demandas repetitivas” (Ibid., p. 100-101).

¹⁴⁷¹ “O atual estado da arte da evolução tecnológica já permite o desenvolvimento de um sistema computacional programado para elaborar, sem intervenção humana, uma minuta de sentença judicial válida e adequadamente fundamentada conforme os parâmetros estabelecidos no art. 489 do Código de Processo Civil Brasileiro a partir de um determinado caso concreto, não sendo necessário o desenvolvimento de uma Inteligência Artificial forte para a realização desta tarefa, bastando à utilização de técnicas informáticas e de gestão de conhecimento já existentes e acessíveis” (VALENTINI, Rômulo Soares. Op. cit., p. 137).

¹⁴⁷² “As noted above, whilst AI has the potential to replace current human judicial functions in terms of some aspects of adjudicative work, technological advances are more likely to support human judges in their judicial work. In this regard, a number of commentators have noted that the goal of the development of AI systems should be to complement current human work, allowing for greater efficiencies, rather than total replacement of humans. At times, these developments suggest that ‘co-bots’ rather than robots will play a more important role in Judge AI. AI programs that can produce a decision based on information input could be used to assist human judges, rather than replace them. These systems could produce a draft judgment based on the system’s determined outcome. A human judge could then use this draft judgment to produce their own reasons, allowing for human oversight over the computer program, and enabling discretionary or social considerations to be made that may be beyond the capacity of the computer program” (SOURDIN, Tania. Judge v Robot? Artificial Intelligence and Judicial Decision-Making. **UNSW Law Journal**, Sydney, v. 41, n. 4. p. 1114-1133, 2018).

pelos juízes, uma vez que a inteligência humana tem características que devem ser mantidas ou são difíceis de replicar em sistemas computacionais.¹⁴⁷³

Por sua vez, Lance Eliot é um crítico da utilização dos termos “juiz-robô” (“*robot judge*”) e “advogado robô” (“*robot lawyer*”)¹⁴⁷⁴, pois, numa análise que transcende o posicionamento que considera um purismo simplório¹⁴⁷⁵, o autor entende que a simplificação do problema por meio de um conceito de “juiz-robô”¹⁴⁷⁶ passa a equivocada e confusa ideia de uma necessária eliminação do elemento humano no Direito.¹⁴⁷⁷

Todavia, no contexto da presente obra, conforme foi exposto no subcapítulo 3.2, o **robô processual** reflete a possibilidade de automação parcial dos andamentos processuais, especificamente em relação às atividades que independem de decisão judicial ou que são relacionadas aos atos de maior simplicidade, decorrentes de despachos judiciais sem caráter decisório. Destaca-se, por outro lado, a maior amplitude e o dinamismo do conceito de **juiz ciborgue** desenvolvido nesta obra, cuja racionalidade é aprimorada pelo duplo canal de correção.

Desse modo, o paradigma de robô processual depende de conhecimento em desenvolvimento de *software*, o que, por padrão, pode ser realizado por equipes centralizadas dos Tribunais. Contudo, a centralização do desenvolvimento também

¹⁴⁷³ “Any shift towards Judge AI clearly raises many issues which have so far been the subject of limited commentary. Some initial issues relate to the exercise of discretion and to what extent judges should – not could – be replaced by AI. It is also important to understand that judges do far more than make decisions. They manage cases, provide a responsive and human framework, settle cases, manage court systems and processes as well as playing an important public and educative function. Drawing the boundaries of acceptable Judge AI requires consideration of ethical questions, as well as questions about who produces algorithms and Judge AI and the extent to which discretion and oversight will be maintained within the judiciary.”

In addition, it may be unhelpful to maintain that Judge AI will only ever stand apart in some way from judges. As noted above, many technology futurists suggest that it is likely that humans will not necessarily be replaced by AI. Instead, human intelligence is likely to be supplemented by technological advances. This approach suggests that judges may remain human but be ‘supplemented’; that is, have their intelligence and analytical functions supported by AI. These enhancement approaches raise issues about judicial appointment, workload and retention and broader questions about how judges contribute to society, as well as the importance of responsive judging and a need to better understand and explore the impact that people experience when a human judge deals with their concerns” (Id.).

¹⁴⁷⁴ “Perhaps we’ll go into courtrooms one day and see a robot judge (robo-judges) and robot lawyers, all awaiting those irascible humans to come in to legal crucible to seek justice for human indiscretions against fellow humans” (ELIOT, Lance. Op. cit., p. 23).

¹⁴⁷⁵ “As such, purists would assert that we ought to ban the use of the robot lawyer moniker altogether” (Ibid., p. 24).

¹⁴⁷⁶ “Stop the use of it, curtail the click-bait, and hopefully bury, until or if we someday really to achieve an AI capable of legal reasoning and can do the same as human lawyers” (Ibid., p. 25).

¹⁴⁷⁷ “Henceforth herein, I’m going to refrain from referring to robo-lawyer or robot lawyer, since I too find the phrasing to be misleading and confounding” (Ibid., p. 27).

tem desvantagens (vide subcapítulo 3.2), e a principal delas é o tolhimento às novas abordagens que possam surgir à margem da anuência institucional da Administração dos Tribunais e do Conselho Nacional de Justiça. Esse é um ponto de tensão que tem origem na necessária prevenção no desenvolvimento de sistemas de IA judicial, o que não deve ser visto como uma barreira ao desenvolvimento científico, pois essa questão deve ser abordada pela administração judiciária com um enfoque de abertura às iniciativas tecnológicas nascidas na prática jurisdicional.

Conforme já anteriormente destacado, em uma abordagem de design dirigido ao domínio¹⁴⁷⁸, a regulação centralizadora e fossilizadora do desenvolvimento de *software* judicial pode representar barreira ao avanço tecnológico e ao progresso da ciência, pois tem como premissa a ausência de comunicação entre a equipe de desenvolvimento do sistema e o julgador que utilizará o *software* nas rotinas judiciais (modelo de automação parcial dos andamentos processuais), o que, além de desconsiderar as vantagens do modelo *domain-driven design*, não percebe os demais problemas jurídicos que isso acarreta. Esse encastelamento é prejudicial. A cautela é necessária para evitar abusos (especialmente através da pura delegação de decisões judiciais à Inteligência Artificial), mas a regulação não pode ser fruto de um medo paralisante que impeça o desenvolvimento científico.

O art. 20 da Lei Geral de Proteção de Dados (Lei 13.709/2018) autoriza a simples automatização, desde que resguardado o direito à revisão humana. Já a interpretação sistemática da Resolução 332/2020 do CNJ demonstra que a automação¹⁴⁷⁹ processual é permitida. Nesse sentido, embora o *caput* do art. 23 da Lei 13.709/2018 sugira cautela para a utilização de modelos de Inteligência Artificial em matéria penal, no §1º do mesmo artigo é plenamente permitido o desenvolvimento de sistemas computacionais voltados à “automação” e ao apoio para cálculo de penas, contagem de prescrição, análise de reincidência e para “mapeamentos, classificações e triagem dos autos para fins de gerenciamento de acervo”.¹⁴⁸⁰

¹⁴⁷⁸ “Ouça a linguagem utilizada pelos especialistas do domínio. Existem termos que sucintamente expressam algo complicado? Eles estão corrigindo sua escolha de palavras (talvez diplomaticamente)? Os olhares confusos em suas faces somem quando você usa determinada frase? Esses são indícios de um conceito que pode ajudar o modelo” (EVANS, Eric. Op. cit., p. 198).

¹⁴⁷⁹ No contexto da presente tese a automação, desde que realizada na busca de se atribuir grau de autonomia decisória ao agente automatizado, é considerada inteligência artificial, conforme exposto no subcapítulo 3.2.

¹⁴⁸⁰ CONSELHO NACIONAL DE JUSTIÇA. **Resolução nº 332, de 21 de agosto de 2020.**

Desse modo, o trâmite do processo judicial pode ser objeto de maior grau de automatização, devendo ser guardada uma maior cautela para o modelo hibridizado de decisão humano-máquina do juiz ciborgue.

5.2 A Obtenção das Bases Dados (*Datasets*) para o Desenvolvimento de IA Judicial

Na trilha rumo ao desenvolvimento de Inteligência Artificial aplicada ao Poder Judiciário, a obtenção de uma base de dados adequada¹⁴⁸¹ é um passo imprescindível para o treinamento de modelos de Inteligência Artificial baseados em aprendizado de máquina.

Os dados podem ser obtidos através do Datajud – Base Nacional de Dados do Poder Judiciário¹⁴⁸², cuja previsão normativa encontra-se na Resolução 331/2020 do Conselho Nacional de Justiça¹⁴⁸³. O Datajud é uma base de dados centralizada que deve concentrar dados e metadados¹⁴⁸⁴ dos processos judiciais da maioria dos Tribunais brasileiros.¹⁴⁸⁵

¹⁴⁸¹ “*The first step is to collect and process the raw data, a corpus of natural language legal texts*” (ASHLEY, Kevin D. **Artificial Intelligence and Legal Analytics**, p. 236).

¹⁴⁸² CONSELHO NACIONAL DE JUSTIÇA. **Datajud**: Base Nacional de Dados do Poder Judiciário. [2022]. Disponível em: <https://www.cnj.jus.br/sistemas/datajud/>. Acesso em: 12 set. 2022.

¹⁴⁸³ CONSELHO NACIONAL DE JUSTIÇA. **Resolução nº 331, de 20 de agosto de 2020**. Institui a Base Nacional de Dados do Poder Judiciário – DataJud como fonte primária de dados do Sistema de Estatística do Poder Judiciário – SIESPJ para os tribunais indicados nos incisos II a VII do art. 92 da Constituição Federal. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3428>. Acesso em: 12 set. 2022.

¹⁴⁸⁴ “Art. 2º Para os fins desta Resolução, considera-se:

I – Metadados processuais: informações estruturadas dos processos judiciais;

II – Serviço de dados: solução que possibilita a interação e integração entre aplicações, permitindo a comunicação de dados e a interoperabilidade entre sistemas desenvolvidos em plataformas diferentes;

III – Período de referência: marco temporal que indica o período ao qual os dados remetidos se referem;

IV – API: é a abreviação de “Application Programming Interface”, em vernáculo, “Interface de Programação de Aplicativos”, e corresponde a um conjunto de instruções e padrões de sistema que possibilitam integração e intercâmbio de dados.

Art. 3º O DataJud será alimentado com dados e metadados processuais relativos a todos os processos físicos ou eletrônicos, públicos ou sigilosos, de qualquer das classes previstas nas Tabelas Processuais Unificadas – TPUs, criadas pela Resolução CNJ nº 46/2007” (Id.).

¹⁴⁸⁵ Atualmente o Supremo Tribunal Federal possui precedentes no sentido de que não se submete ao controle administrativo e à gestão do Conselho Nacional Justiça, motivo pelo qual a Resolução 331/2020 infelizmente não se aplica à Corte Constitucional (conforme recorte adotado pela própria normativa), solução que este autor entende inadequada como política de tecnologia da informação e do ponto de vista constitucional. Deve existir distinção entre a impossibilidade do CNJ realizar a análise disciplinar de Ministros do STF (o que se harmoniza com o art. 52, II, da Constituição Federal) e a possibilidade do CNJ regular boas práticas administrativas em relação ao mesmo Tribunal (art. 103-B, §4º, I e II, da Constituição Federal). Todavia, o debate extrapola os limites da presente obra e não cabe, em toda sua amplitude, na presente nota de rodapé.

Conforme previsto no art. 11, *caput*, da Resolução 331/2020 do CNJ¹⁴⁸⁶, desde que respeitados o sigilo e a confidencialidade das informações (na forma da Lei 13.709/2018), o acesso aos metadados processuais pode ser realizado através de uma API pública¹⁴⁸⁷, mediante controle prévio do Conselho Nacional de Justiça.¹⁴⁸⁸ Já o acesso aos dados processuais depende de requerimento formal de ente público ou de instituição de pesquisa interessada (art. 11, Parágrafo único, da Resolução 331/2020 do CNJ¹⁴⁸⁹), porém a instituição interessada no acesso aos dados deverá firmar termo específico de sigilo e confidencialidade junto ao CNJ.

Com o início da implementação no ano de 2022, o Conselho Nacional de Justiça lançou a plataforma Codex (Resolução 446/2022 do CNJ¹⁴⁹⁰), a qual se volta à disponibilização de acesso a “dados, metadados processuais e inteiro teor dos documentos e atos proferidos relativos a todos os processos eletrônicos, públicos ou sigilosos” (art. 2º, *caput*, da Resolução 446/2022 do CNJ).¹⁴⁹¹

Assim, os Tribunais têm acesso ao Datajud e à plataforma Codex, o que lhes possibilita o desenvolvimento de sistemas de Inteligência Artificial através de rica base de dados que será disponibilizada pelo Judiciário brasileiro.

¹⁴⁸⁶ CONSELHO NACIONAL DE JUSTIÇA. **Resolução nº 331, de 20 de agosto de 2020.**

¹⁴⁸⁷ O que ainda pende de maior regulamentação pela Presidência do Conselho Nacional de Justiça.

¹⁴⁸⁸ “Art. 1º O acesso das equipes técnicas dos órgãos do Poder Judiciário e do próprio Conselho Nacional de Justiça à Base Nacional de Dados do Poder Judiciário (DataJud) será concedido para atividades específicas de sustentação tecnológica, de colaboração efetiva no desenvolvimento de projetos, realização de pesquisas, desenvolvimento de ferramentas de disponibilização de dados ou acesso a insumos e dados para projetos de interesse do CNJ ou da presidência do órgão” (CONSELHO NACIONAL DE JUSTIÇA. **Portaria nº 196, de 10 de outubro de 2022.** Institui regras de acesso das equipes técnicas ao DataJud. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3428>. Acesso em: 12 set. 2022).

¹⁴⁸⁹ CONSELHO NACIONAL DE JUSTIÇA. **Resolução nº 331, de 20 de agosto de 2020.**

¹⁴⁹⁰ CONSELHO NACIONAL DE JUSTIÇA. **Resolução nº 446, de 14 de março de 2022.** Institui a plataforma Codex como ferramenta oficial de extração de dados estruturados e não estruturados dos processos judiciais eletrônicos em tramitação no Poder Judiciário Nacional e dá outras providências. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/4417>. Acesso em: 12 set. 2022.

¹⁴⁹¹ “Art. 2º O Codex é alimentado com dados, metadados processuais e inteiro teor dos documentos e atos proferidos relativos a todos os processos eletrônicos, públicos ou sigilosos, de quaisquer das classes previstas nas Tabelas Processuais Unificadas (TPUs), criadas pela Resolução CNJ nº 46/2007, e que estejam em tramitação no Poder Judiciário ou tenham sido baixados a partir de 1º de janeiro de 2020.

§ 1º Os metadados extraídos do Codex representarão, da melhor forma, os dados armazenados em sua origem.

§ 2º Compete ao tribunal realizar, sempre que necessário, o processo de transformação e saneamento desses dados, em linha com estabelecido na Resolução CNJ nº 331/2020, de forma que as informações recebidas no CNJ estejam em consonância com as Classes, Assuntos, Movimentos e Documentos das Tabelas Processuais Unificadas, com os códigos dos órgãos julgadores cadastrados no Sistema Corporativo do CNJ, bem como realizar os demais procedimentos de limpeza e transformação de dados” (Id.).

Esse é um lado da moeda, pautado na centralização dos dados e no controle de acesso, o que é necessário para garantir a integridade e a segurança da informação.

Todavia, existem ainda outras formas de acesso às bases de dados. Eles podem ser obtidos diretamente junto aos Tribunais, haja vista a autonomia desses para gerir seus próprios sistemas de Inteligência Artificial judicial (art. 98, I, “b”, e art. 99, *caput*, da Constituição Federal).

Além disso, em respeito à independência judicial, caso haja dificuldade da equipe técnica da unidade judicial na obtenção de dados (pois o julgador deve ser um dos protagonistas no processo de desenvolvimento¹⁴⁹²), soluções mais maleáveis e com maior resiliência (dentro dos limites legais e constitucionais) podem ser adotadas. Pois a decisão judicial não pode sofrer a limitação de terceiros no aspecto aqui delineado, razão pela qual o Juiz e a sua equipe podem buscar os dados, tratar os dados necessários ao treinamento de sistemas de Inteligência Artificial baseados em aprendizado de máquina e desenvolver os sistemas de IA judicial.

A partir desse modelo, o julgador pode utilizar as bases de dados públicas a que tem acesso, a fim de que, respaldado pela sua garantia de independência judicial na obtenção de dados voltados à melhor aplicação do Direito no julgamento (o que é parte da pretensão de correção), utilize técnicas de raspagem de dados (*scraping*¹⁴⁹³). Nesse caso, é destacada a importância do assessor-desenvolvedor para a implementação de técnica de raspagem de dados¹⁴⁹⁴ e para o pré-processamento da base de dados (*dataset*).

5.3 Data Science e Machine Learning: Uma Visão Necessária do Aprendizado de Máquina voltado à IA Judicial

É comum pensar que Direito é também uma arte, não apenas uma técnica. Guardadas as devidas proporções, tendo em vista que os modelos de IA baseados em análise de dados necessitam de mineração de dados, bem como, levando em

¹⁴⁹² E a figura do assessor-desenvolvedor, anteriormente tratada nesta obra, será necessária.

¹⁴⁹³ DIOUF, Rabiyou; SARR, Edouard Ngor; SALL, Ousmane; BIRREGAH, Babiga; BOUSSO, Mamadou; MBAYE, Sény Ndiaye. Web Scraping: State-of-the-Art and Areas of Application. **2019 IEEE International Conference on Big Data (Big Data)**, Los Angeles, p. 6040-6042, 2019. DOI: 10.1109/BigData47090.2019.9005594.

¹⁴⁹⁴ Por exemplo, o assessor-desenvolvedor pode utilizar a biblioteca *Selenium*, disponível para as linguagens Python e Java, dentre outras tecnologias existentes.

consideração que não existe um único paradigma de IA que seja a “mágica” por trás das soluções, pode-se falar, em certos aspectos, que a “mineração de dados é uma arte”¹⁴⁹⁵ que impacta na modelagem da IA judicial.

Para que sistemas baseados em aprendizado de máquina sejam construídos, é necessário proceder à abordagem típica da Ciência de Dados. Desse modo, é preciso, antes de tudo, esquadrihar o problema relacionado à decisão judicial, a fim de que a mineração de dados jurídicos leve à extração de conhecimento útil¹⁴⁹⁶ à aplicação do Direito, uma vez que a partir dos dados jurídicos a Ciência de Dados pode encontrar atributos descritivos¹⁴⁹⁷ de como determinados problemas jurídicos se manifestam.

Para tanto, a tomada de decisão orientada por dados (*data-oriented design – DOD*)¹⁴⁹⁸ serve como auxílio ao julgador, o que leva à possibilidade de “descoberta” de conhecimento a partir dos dados e de mapeamento de decisões que se repetem¹⁴⁹⁹, destacando que neste modelo, “a tomada de decisão pode se beneficiar até mesmo de pequenos aumentos na precisão deste processo com base em análise de dados”.¹⁵⁰⁰ Dessa maneira, caso adequadamente implementados, os modelos preditivos de IA podem abstrair parte da complexidade¹⁵⁰¹ do mundo jurídico através de atributos (*features*) que levem à percepção do suporte fático concreto e/ou dos tipos de argumentos utilizados.

¹⁴⁹⁵ “A mineração de dados é uma arte. Ela envolve a aplicação de uma quantidade substancial de ciência e tecnologia, mas a aplicação adequada ainda envolve arte também. Mas, como acontece em muitas artes maduras, existe um processo bem compreendido que coloca uma estrutura no problema, permitindo consistência, repetitividade e objetividade razoáveis. Uma codificação útil do processo de mineração de dados é dada pelo Processo Padrão da Indústria Cruzada para Exploração de Dados” (PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 27).

¹⁴⁹⁶ “Conceito fundamental: extrair conhecimento útil a partir de dados para resolver problemas de negócios pode ser tratado de forma sistemática, seguindo um processo com etapas razoavelmente definidas” (Ibid., p. 14).

¹⁴⁹⁷ “Conceito fundamental: a partir de uma grande massa de dados, a tecnologia da informação pode ser usada para encontrar atributos descritivos informativos de entidades de interesse” (Id.).

¹⁴⁹⁸ “Tomada de decisão orientada por dados (DOD) refere-se à prática de basear as decisões na análise de dados, em vez de apenas na intuição. Por exemplo, um negociante poderá selecionar anúncios baseados puramente em sua longa experiência na área e em sua intuição de que funcionará. Ou, pode basear sua escolha na análise dos sobre a forma como os consumidores reagem a diferentes anúncios. Ele também pode utilizar uma combinação dessas abordagens. A DOD não é uma prática do tipo ‘tudo ou nada’” (Ibid., p. 5).

¹⁴⁹⁹ Ibid., p. 6.

¹⁵⁰⁰ Id.

¹⁵⁰¹ “No momento, é suficiente entender que um modelo preditivo abstrai a maior parte da complexidade do mundo, concentrando-se como um conjunto específico de indicadores que se correlacionam, de algum modo, com uma quantidade de interesses (quem apresentará rotatividade ou quem comprará, quem está grávida, etc)” (Id.).

Além disso, tendo em vista que nem todos os dados podem estar previamente disponíveis ou devidamente estruturados¹⁵⁰², a obtenção dos dados para os sistemas de Inteligência Artificial judicial pode envolver custos¹⁵⁰³, o que deve ser levado em conta no momento da definição de quais dados serão explorados¹⁵⁰⁴ e qual é a potencial capacidade de extrair conhecimento útil a partir deles.¹⁵⁰⁵

Nessa linha, o desenvolvimento de Inteligência Artificial jurídica demanda que o profissional do Direito conheça as ferramentas de Ciência de Dados, uma vez que a atual fase de desenvolvimento do campo de estudo é a implementação do Direito aliado às ferramentas de Ciência de Dados, o que guarda a estrita relação entre teoria e prática experimental.¹⁵⁰⁶ A equipe de implementação (composta por desenvolvedores, profissionais do Direito¹⁵⁰⁷ e por assessores-desenvolvedores) deve interagir¹⁵⁰⁸, pois, em cada projeto de desenvolvimento de sistema de IA judicial, a experiência da equipe interdisciplinar (e do assessor-desenvolvedor) é necessária para a melhor modelagem a partir dos dados, da seleção dos atributos (*features*) a serem objeto do treinamento e para evitar o sobreajuste (*overfitting*).¹⁵⁰⁹

¹⁵⁰² Nem sempre a base de dados adequada existirá ou estará disponível nos repositórios oficiais, o que demandará bastante trabalho na busca dos dados e na organização da base de dados.

¹⁵⁰³ (...) “uma estratégia fundamental de data science: obter os dados necessários a um custo. Depois que visualizamos dados como um ativo de negócios, devemos pensar se e quanto estamos dispostos a investir.” (...) “O pensador analítico de dados precisa considerar se espera que os dados tenham valor suficiente para justificar o investimento” (PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 10).

¹⁵⁰⁴ TAULLI, Tom. Op. cit., p. 39-41.

¹⁵⁰⁵ (...) “um dos princípios fundamentais de data science: os *dados*, e a *capacidade de extrair conhecimento útil a partir deles*, devem ser considerados importantes ativos estratégicos” (PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 9).

¹⁵⁰⁶ “O estado do data science pode ser comparado ao da química em meados do século XIX, quando teorias e princípios gerais estavam sendo formulados e o campo era, basicamente experimental. Todo bom químico tinha que ser um técnico de laboratório competente. Da mesma forma, é difícil imaginar um cientista de dados que não está acostumado com certos tipos de ferramentas de software” (Ibid., p. 16).

¹⁵⁰⁷ “Uma verdade é insofismável, julgar não é fazer meras análises literais e objetivas da lei. Há que se observar questões éticas e interpretar a subjetividade de cada caso. Um raciocínio que as máquinas não são capazes de conduzir. A sobrevida humana no comando das Cortes ainda terá lugar” (LAGE, Fernanda de Carvalho. Op. cit., p. 110).

¹⁵⁰⁸ “Em uma escala menos grandiosa, mas provavelmente mais comum, projetos de análise de dados alcançam todas as unidades de negócios. Os funcionários em todas essas unidades devem interagir com a equipe de data science. Se esses funcionários não têm uma base fundamental nos princípios de pensamento analítico de dados, eles não vão realmente entender o que está acontecendo na empresa. Esta falta de compreensão é muito mais prejudicial em projetos de data science do que em outros projetos técnicos, pois data science dá suporte a uma tomada de decisão melhorada” (PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 13).

¹⁵⁰⁹ “Conceito fundamental: se olhar muito para um conjunto de dados, encontrará alguma coisa –, mas isto pode não ser generalizável além dos dados para os quais está olhando. Isto é chamado de sobreajustar um conjunto de dados. Técnicas de mineração de dados podem ser muito poderosas, e a necessidade de detectar e evitar sobreajuste é um dos conceitos mais importantes para se compreender quando se aplica mineração de dados para problemas reais. O conceito de sobreajuste

Assim, a mineração de dados jurídicos pode auxiliar a tomada de decisões judiciais¹⁵¹⁰, uma vez que a utilização de dados para a modelagem de soluções de Inteligência Artificial judicial necessita que se pense com cuidado a respeito do contexto de utilização (o problema a ser resolvido e a forma de mensurar a resolução do mesmo¹⁵¹¹) para que, a partir dos dados (quais dados podem ser utilizados como atributos para treinamento do modelo?), seja definido o que é útil para a resposta ao problema.¹⁵¹²

Desse modo, a grande disponibilidade de dados possibilita a implementação de modelos de IA que se aproveitam do modelo de decisão baseada em dados, a fim de extrair informações¹⁵¹³ e conhecimento úteis¹⁵¹⁴ à tomada da decisão judicial baseada no duplo canal de correção.

5.4 Ciência de Dados e Direito: Principais Aplicações de Modelos de IA Judicial Baseados em Aprendizado de Máquina

Tendo por base as premissas anteriormente discutidas, passa-se à necessária exposição das principais espécies de modelos de Inteligência Artificial que podem ser utilizadas no desenvolvimento de sistemas inteligentes de apoio ao Poder Judiciário.

A **classificação e a estimativa de probabilidade de classe** “tentam prever, para cada indivíduo de uma população, a que (pequeno) conjunto de classes este indivíduo pertence”¹⁵¹⁵, o que o faz através de uma pontuação ou estimativa de

e sua prevenção permeia os processos de data science, de algoritmos e de métodos de avaliação” (Ibid., p. 15).

¹⁵¹⁰ “Data science suporta a tomada de decisões orientada por dados – e, às vezes, conduz automaticamente tais tomadas de decisão – e depende de tecnologias para armazenamento e engenharia de ‘Big Data’, mas seus princípios são separados” (Ibid., p. 17).

¹⁵¹¹ TAULLI, Tom. Op. cit., p. 50.

¹⁵¹² “Conceito fundamental: *formular soluções de mineração de dados e avaliar os resultados envolve pensar cuidadosamente sobre o contexto em que serão utilizados*. Se nosso objetivo é a extração de conhecimento potencialmente útil, como podemos formular o que é útil? Isso depende criticamente da aplicação em questão” (PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 15).

¹⁵¹³ “*Applying ML to legal texts will play key roles in a legal app for cognitive computing. One goal in cognitive computing is for ML algorithms to learn to identify* Applying ML to legal texts will play key roles in a legal app for cognitive computing. One goal in cognitive computing is for ML algorithms to learn to identify” (ASHLEY, Kevin D. **Artificial Intelligence and Legal Analytics**, p. 235).

¹⁵¹⁴ “Essa ampla disponibilidade de dados levou ao aumento do interesse em métodos para extrair informações úteis e conhecimento a partir de dados – o domínio de data science” (PROVOST, Foster; FAWCETT, Tom. Op. cit., p.1).

¹⁵¹⁵ Ibid., p. 20.

probabilidade para a classe.¹⁵¹⁶ No âmbito jurídico, pode-se exemplificar algumas aplicações a esses modelos classificatórios:

a) classificar argumentos nas petições ou nas decisões: trata-se de um argumento de precedente¹⁵¹⁷ ou de um argumento genético, por exemplo?

b) classificar documentos num processo: o documento é uma petição, uma procuração ou uma prova documental? Se é uma petição, o documento é uma petição de juntada de custas ou um recurso? Se é um recurso, é uma apelação ou são embargos de declaração?

c) classificar os casos a partir do suporte fático concreto;

d) classificar o aspecto fático de provas produzidas (testemunhos, depoimentos, perícias, documentos, etc.).

Aqui o desafio do assessor-desenvolvedor é a compreensão do contexto, pois, por exemplo, a busca para classificar argumentos de precedente leva em conta características necessárias para que o sistema perceba o mesmo, o que se faz através do tipo e disposição do texto, da referência ao tribunal julgador, de determinados argumentos utilizados, etc. Em resumo, o treinamento de modelos de classificação utilizará atributos (*features*) da base de dados que levem a inferir o resultado.

Por exemplo, a partir da visão de um profissional do Direito, pode-se pensar em quais são os elementos que diferenciam uma emenda à inicial de um recurso de apelação: a) momento da petição; b) pedido realizado; c) temas abordados no texto; d) endereçamento; e e) outros elementos. Nesse caminho, podem ser selecionados, levando-se em conta a base de dados, quais os atributos que produzem os melhores resultados no treinamento de modelos de IA e os modelos que são os melhores em termos de acurácia, sensibilidade, precisão e outros parâmetros.

Em outro exemplo, a partir de um determinado suporte fático concreto (uma negativa de fornecimento de um medicamento específico em ação ajuizada contra o Estado), um modelo de IA judicial pode ser treinado através de atributos (*features*) para que classifique casos em que determinadas características de saúde da parte são verificadas, a fim de auxiliar na análise do caso, de maneira a melhorar a celeridade nas situações mais evidentes de caracterização do direito ao medicamento

¹⁵¹⁶ “O modelo de pontuação aplicado a um indivíduo produz, em vez de uma previsão de classe, uma pontuação que representa a probabilidade (ou outra quantificação de probabilidade) de que o indivíduo pertença a cada classe” (Ibid., p. 20-21).

¹⁵¹⁷ ALEXY, Robert. **Teoria da Argumentação Jurídica**, p. 240-241.

ou aos mais urgentes em termos de risco à saúde. Dessa forma, sem retirar a decisão final da pessoa do magistrado, tratar-se-ia de um incremento do conhecimento dos elementos fáticos e normativos para o processo de tomada de decisão.

Já a **regressão** (estimativa de valor), “tenta estimar ou prever, para cada indivíduo, o valor numérico de alguma variável”¹⁵¹⁸, sendo que um procedimento de regressão resulta num modelo que calcula o valor da variável específica para o indivíduo.¹⁵¹⁹ Nesse sentido, enquanto a classificação busca prever **se** alguma coisa aconteceu (existe algo?), por sua vez a regressão busca prever **quanto** de algo vai acontecer.¹⁵²⁰ Por exemplo, modelos de regressão podem ser utilizados para investigar:

- a) a extensão e a proporcionalidade de uma indenização civil¹⁵²¹;
- b) a proporcionalidade de uma multa contratual¹⁵²²;
- c) a proporcionalidade de uma multa processual por descumprimento de obrigação de tutela específica.¹⁵²³

Por sua vez, a **combinação por similaridade** “tenta identificar indivíduos semelhantes com base nos dados conhecidos sobre eles”¹⁵²⁴, sendo que as medidas de similaridade baseiam outras tarefas de mineração de dados, como a classificação, a regressão e o agrupamento.¹⁵²⁵ Desse modo, a combinação por similaridade pode ser útil para:

- a) análise de precedentes;

¹⁵¹⁸ PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 21.

¹⁵¹⁹ Id.

¹⁵²⁰ Id.

¹⁵²¹ “Art. 944. A indenização mede-se pela extensão do dano.

Parágrafo único. Se houver excessiva desproporção entre a gravidade da culpa e o dano, poderá o juiz reduzir, eqüitativamente, a indenização” (BRASIL. **Lei 10.406, de 10 de janeiro de 2002**. Código Civil. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/2002/l10406compilada.htm. Acesso em: 27 nov. 2022).

¹⁵²² “Art. 413. A penalidade deve ser reduzida eqüitativamente pelo juiz se a obrigação principal tiver sido cumprida em parte, ou se o montante da penalidade for manifestamente excessivo, tendo-se em vista a natureza e a finalidade do negócio” (Id.)

¹⁵²³ “Art. 537. A multa independe de requerimento da parte e poderá ser aplicada na fase de conhecimento, em tutela provisória ou na sentença, ou na fase de execução, desde que seja suficiente e compatível com a obrigação e que se determine prazo razoável para cumprimento do preceito. § 1º O juiz poderá, de ofício ou a requerimento, modificar o valor ou a periodicidade da multa vincenda ou excluí-la, caso verifique que:

I - se tornou insuficiente ou excessiva;

II - o obrigado demonstrou cumprimento parcial superveniente da obrigação ou justa causa para o descumprimento” (BRASIL. **Lei 13.105, de 16 de março de 2015**).

¹⁵²⁴ PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 21.

¹⁵²⁵ Id.

b) análise de reiteração de ação já anteriormente ajuizada (prevenção, litispendência e coisa julgada);

c) análise de lides multitudinárias;

d) análise de argumento baseado em analogia.

O **agrupamento** (ou clusterização¹⁵²⁶) busca “reunir indivíduos de uma população por meio de sua similaridade, mas não motivado por um propósito específico”.¹⁵²⁷ Isso torna o agrupamento útil para a exploração inicial dos dados (desde a fase de pré-processamento), a fim de identificar grupos naturalmente existentes¹⁵²⁸, que podem inferir novas tarefas de mineração de dados.¹⁵²⁹ Alguns exemplos de usos do agrupamento são:

a) agrupar decisões semelhantes a partir do resultado do julgamento em relação a determinado suporte fático concreto;

b) agrupar decisões semelhantes a partir dos argumentos utilizados para se fundamentar determinado resultado do julgamento;

c) agrupar casos semelhantes a partir do suporte fático concreto;

d) agrupar processos semelhantes a partir dos argumentos utilizados pelas partes (argumentos jurídicos ou práticos do tipo geral, para além do suporte fático concreto);

e) agrupar processos semelhantes a partir dos argumentos utilizados na fundamentação das decisões;

f) agrupar processos semelhantes a partir das provas produzidas;

g) agrupar processos semelhantes a partir dos andamentos, da ordem e das espécies de atos processuais produzidos durante o trâmite processual (o que é útil para inferir cerceamento de defesa, por exemplo).

¹⁵²⁶ “Problemas de Agrupamento também são conhecidos como problemas de clusterização, e seu objetivo é separar os registros de um conjunto de dados em subconjuntos ou grupos (clusters), de tal forma que elementos em um cluster compartilhem um conjunto de propriedades comuns que os diferencie dos elementos de outros clusters. São os problemas de aprendizagem não supervisionada mais comuns” (ESCOVEDO, Tatiana; KOSHIYAMA, Adriano. **Introdução a Data Science: Algoritmos de Machine Learning e métodos de análise**. São Paulo: Casa do Código, 2020, posição 3030. E-book Kindle).

¹⁵²⁷ PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 21.

¹⁵²⁸ “*Document clustering algorithms, for instance, do not use labeled training data. They group instances together without any preexisting information about what labels are correct (see Kohavi and Provost, 1998). Instead, they infer groups of documents based on their content or metadata and leave for humans the task of determining post hoc what the members of the group share and what labels to apply, if any*” (ASHLEY, Kevin D. **Artificial Intelligence and Legal Analytics**, p. 247).

¹⁵²⁹ PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 21.

Para além disso, o agrupamento também pode ser usado como *input*¹⁵³⁰ de auxílio à tomada de decisão judicial com enfoque em perguntas como:

- a) quais classes de argumentos melhor representam a solução para o caso?
- b) quais argumentos específicos (a partir de um banco de soluções¹⁵³¹) melhor representam a solução para o caso (suporte fático concreto)?

Noutro aspecto, o **agrupamento por coocorrência**¹⁵³² busca por associações entre entidades com base em transações que as envolvem.¹⁵³³ O agrupamento por coocorrência resulta numa descrição dos itens que ocorrem conjuntamente (o que pode incluir estatísticas sobre a frequência da coocorrência).¹⁵³⁴ Seguem alguns exemplos de perguntas de coocorrência:

- a) quais fatos comumente ocorrem juntos¹⁵³⁵ para determinada solução?
- b) quais argumentos são utilizados juntos para determinada solução?
- c) quais andamentos processuais geralmente ocorrem para que haja uma determinada solução?
- d) quais provas coocorrem quando há certa solução?

Destarte, o agrupamento de coocorrência leva em conta a similaridade dos objetos com base nas aparições conjuntas de seus elementos.¹⁵³⁶ Essa espécie de solução em mineração de dados é tipicamente utilizada em sistema de recomendação baseado em agrupamento por afinidade (por exemplo, recomendação de músicas, livros, etc.).¹⁵³⁷

Por sua vez, **perfilamento**, que também é conhecido como descrição de comportamento, tem a pretensão de traçar os elementos descritivos típicos do comportamento de indivíduos, grupos ou populações.¹⁵³⁸ Destaca-se que o comportamento, muitas vezes, não tem uma descrição simples, pois pode demandar que se trace um perfil alvo a partir de diversas variáveis¹⁵³⁹ (por exemplo, para se

¹⁵³⁰ Id.

¹⁵³¹ E aqui se manifesta uma faceta do modelo *case based reasoning*.

¹⁵³² Que também é chamado de mineração de conjunto de itens frequentes, de descoberta da regra de associação e de análise de portfólio de ações.

¹⁵³³ PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 21-22.

¹⁵³⁴ Ibid., p. 22.

¹⁵³⁵ E aqui há uma ligação entre suporte fático concreto e suporte fático abstrato desde a concepção da aplicação da IA judicial. A abordagem a partir de fatores (*factors*) é uma inspiração natural à abordagem.

¹⁵³⁶ PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 21-22.

¹⁵³⁷ TAULLI, Tom. Op. cit., p. 88.

¹⁵³⁸ PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 22.

¹⁵³⁹ Id.

traçar o perfil de espécies casos ou de tendências de julgadores). Todavia, deve existir cuidado ético com o uso das técnicas de perfilamento, especialmente para não se incorrer em enviesamento racial ou de grupos étnicos, como ocorreu no muito relatado caso do sistema COMPAS.¹⁵⁴⁰

O perfilamento, ainda que lembre, à primeira vista, o estabelecimento de perfis pessoais a serem buscados (o que é problemático em julgamentos criminais, embora possa ter utilidade prática em investigações policiais), pode ser utilizado para pré-análise de processos ou de decisões em busca de características que estabeleçam um perfil de casos a serem recomendadas determinadas soluções. No mais, em um passo mais ligado à argumentação jurídica, Kevin Ashley descreve técnicas de perfilamento para o estabelecimento de “perfis semânticos dos estatutos legais” (*“semantic profiles of statutes”*).¹⁵⁴¹

Para além disso, o perfilamento pode ser destinado à verificação do comportamento de agentes, o que tem especial utilidade na detecção de anomalias e fraudes, apresentando aplicabilidade para a verificação de invasões de sistemas computacionais (pode ser utilizado como técnica auxiliar na produção da prova técnica, convertendo-se ele mesmo em um argumento do típico prático) ou fraudes contra consumidores (a partir de compras que um consumidor normalmente faz ou não). Nesse sentido, pode-se “usar o **grau de disparidade** como uma pontuação suspeita e emitir um alarme, se for muito elevada”.¹⁵⁴²

Desse modo, estas são algumas das possibilidades de utilização de técnicas de perfilamento:

a) medir o grau de disparidade entre decisões de um mesmo julgador a respeito de base fática semelhante (suporte fático concreto);

¹⁵⁴⁰ CHOULDECHOVA, Alexandra. Op. cit.

¹⁵⁴¹ “*Semantic profiles of statutes: Some work combines functional types, abstract categories, or conceptual indexing into a broader semantic profile of statutes or regulations. Wyner and Peters (2011) employed linguistically oriented rules to extract information from a complex, four-page U.S. Food and Drug Administration regulation concerning testing for disease agents in blood. They sought to identify in the text different types of normative rules, including obligations and permissions and their antecedents, subject agents, subject themes, and exceptions. In Winkels and Hoekstra (2012), the authors extracted concepts and definitions from statutory texts related to Dutch tax law, using semantic web technology and NLP techniques*” (ASHLEY, Kevin D. **Artificial Intelligence and Legal Analytics**, p. 262).

¹⁵⁴² PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 22.

b) medir o grau de disparidade entre decisões de um grupo de julgadores a respeito de base fática semelhante (suporte fático concreto) para a análise de casos repetitivos;

c) medir o grau de disparidade entre os argumentos trazidos pela parte e o que foi provado (suporte fático concreto);

d) traçar um perfil de litigantes de má-fé a partir de bases fáticas específicas (casos);

e) traçar um perfil dos casos conforme o suporte fático concreto;

f) traçar um perfil dos casos conforme o suporte fático abstrato;

g) traçar um perfil dos casos conforme os argumentos utilizados (pelas partes ou na fundamentação de decisões sujeitas a recurso).

Na sequência, deve ser destacada a **previsão de vínculo**, técnica que ‘tenta prever ligações entre itens de dados, geralmente sugerindo que um vínculo deveria existir e, possivelmente, também estimando a força do vínculo’.¹⁵⁴³ Por exemplo, a previsão de vínculo é bastante utilizada em sistemas de redes sociais. Dessa maneira, a referida técnica pode ser utilizada, conforme os seguintes exemplos, para:

a) estimar se um precedente tem uma possível solução para o caso ao compartilhar determinadores fatores/elementos com o caso sob julgamento;

b) estimar se deve ser utilizado um determinado argumento na fundamentação a partir do compartilhamento do mesmo suporte fático (concreto, abstrato ou representado no sistema);

c) estimar a força do vínculo através da ponderação do “peso” de cada elemento/fator;

d) a partir de determinada força do vínculo estimado, sugerir qual novo argumento pode ser utilizado a partir do manejo de um primeiro argumento.

Por sua vez, a **redução de dados** busca substituir um grande conjunto de dados por um conjunto de dados menor que mantenha a maior parte das informações importantes do conjunto originário¹⁵⁴⁴, uma vez que, muitas vezes, é mais fácil manipular e processar um conjunto de dados reduzido que pode deixar mais evidenciadas as informações buscadas¹⁵⁴⁵, o que possibilita a utilização de técnicas

¹⁵⁴³ Id.

¹⁵⁴⁴ Ibid., p. 23.

¹⁵⁴⁵ Id.

que não funcionam bem com grandes volumes de dados.¹⁵⁴⁶ Contudo, deve ser destacado que a escolha pelo uso da redução de dados na busca de eficiência computacional acarreta a perda de informação, sendo necessário um equilíbrio no momento de se realizar a redução¹⁵⁴⁷, que só é possível na análise do problema de Ciência de Dados em si (momento em que o aspecto “arte” se manifesta para o refinamento da técnica). Além disso, com a redução de dados pode tornar mais difícil a compreensão do funcionamento do modelo¹⁵⁴⁸, o que deve ser considerado ao se pensar no direito à explicabilidade. Eis algumas das possibilidades de utilização da redução de dados no âmbito jurídico:

a) pré-processamento de dados jurídicos para utilização no treinamento de modelos de aprendizado de máquina (*machine learning*);

b) pré-processamento de grandes séries de dados a serem utilizados como potenciais provas (suporte fático concreto) que embasem a fundamentação pautada em argumentos do tipo prático, a fim de buscar tendências e hipóteses de julgamento (por exemplo, a redução de dados aplicada à análise de documentos contábeis e fiscais);

c) pré-processamento de espécies de argumentos utilizados pelas partes em suas manifestações, a fim de identificar tendências de julgamento;

d) pré-processamento de espécies de argumentos utilizados em decisões de um determinado julgador, a fim de se inferir tendências de decisão.

Também há a **modelagem causal**, com a qual se busca estabelecer um auxílio para a compreensão de quais acontecimentos ou ações efetivamente causam influência em outras pessoas.¹⁵⁴⁹ Destaca-se que as técnicas de modelagem causal podem necessitar de um grande investimento em dados (experimentos randomizados controlados¹⁵⁵⁰ e métodos para alcançar conclusões causais a partir de dados observacionais), sendo que os “métodos experimentais e observacionais para modelagem causal geralmente podem ser visualizados como análises

¹⁵⁴⁶ GRUS, Joel. **Data Science do Zero**. Tradução: Welington Nascimento. Rio de Janeiro: Alta Books, 2016, p. 139.

¹⁵⁴⁷ PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 23.

¹⁵⁴⁸ GRUS, Joel. Op. cit., p. 139.

¹⁵⁴⁹ PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 23.

¹⁵⁵⁰ Por exemplo, através de testes A/B, nos quais são comparados modelo de IA diferentes, na busca de se encontrar o modelo que tenha o desempenho mais eficiente. Para um maior aprofundamento vide: GRUS, Joel. Op. cit., p. 87-88.

‘contrafactuais’¹⁵⁵¹, uma vez que com eles se pretende melhor entender a diferença entre as situações determinadas em que um evento (o qual, por exemplo, pode ser representado pela utilização de um argumento específico) influencia ou não no julgamento. São exemplos da aplicabilidade da modelagem causal:

a) verificar a influência de determinados argumentos utilizados nas decisões de primeiro grau sobre as decisões de grau recursal (a decisão recursal sofreu alteração com base nos argumentos da decisão recorrida ou a decisão recursal ocorreria no mesmo sentido de qualquer maneira?);

b) verificar a influência de determinados fatores (suporte fático concreto) nas decisões de um julgador ou de um grupo de julgadores (a decisão recursal sofreu alteração com base na explicitação de determinado suporte fático concreto, ou a decisão ocorreria no mesmo sentido de qualquer maneira?).

No mais, é uma boa prática incluir a modelagem causal em sistema de IA judicial, a fim de verificar o impacto do funcionamento do sistema em relação à eficiência de outros subsistemas que utilizem IA, para poder se perguntar acerca da eficiência das soluções adotadas ao demandar/desafiar pretensão de correção aos julgadores.

Para além desta visão não exaustiva, porém necessária, de espécies de soluções de IA baseada em Ciência de Dados e suas potencialidades de aplicação para o Direito, outros pontos merecem ser delineados, o que se passa a fazer em seguida.

5.5 *Machine Learning*: Aprendizagem Supervisionada, Aprendizagem Não Supervisionada e Aprendizagem por Reforço

A aplicação do aprendizado de máquina (*machine learning*) para o desenvolvimento de sistemas de Inteligência Artificial voltados ao auxílio do Judiciário se baseia em métodos supervisionados e métodos não supervisionados.¹⁵⁵²

¹⁵⁵¹ PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 23.

¹⁵⁵² “Uma parte vital nas fases iniciais do processo de mineração de dados é (i) decidir se a linha de ataque será supervisionada ou não supervisionada e (ii) caso supervisionada, produzir uma definição precisa de uma variável alvo. Essa variável deve ser uma quantidade específica que será o foco de mineração de dados (e para a qual podemos obter valores para alguns exemplos de dados)” (Ibid., p. 25).

Na **aprendizagem supervisionada** “cada exemplo é associado a uma anotação, uma etiqueta”¹⁵⁵³, de maneira que os algoritmos exploram a base de dados (*dataset*) através dos elementos anotados (etiquetados). Algoritmos de aprendizagem supervisionada são utilizados para tarefas de classificação¹⁵⁵⁴ e de regressão. Nesse paradigma se separa a base de dados em *dataset* de **treino** (que será submetido ao algoritmo de aprendizado de máquina) e *dataset* de **teste** (que são os dados em que o modelo de IA será submetido para avaliação e validação de desempenho). Logo, a obtenção de grandes *datasets*¹⁵⁵⁵ é o segredo para um melhor desenvolvimento de sistemas de Inteligência Artificial judicial.

Desse modo, no paradigma de aprendizado supervisionado há a apresentação de dados conhecidos corretamente rotulados através do *dataset* de treino, a fim de que o algoritmo escolhido¹⁵⁵⁶ possa levar a um modelo de IA que realize a tarefa de classificação ou de regressão com o melhor desempenho possível. Assim, o modelo de IA resulta de um algoritmo que aprende a partir de um conjunto de dados de treinamento.¹⁵⁵⁷

A implementação de IA no Direito através de técnicas de aprendizado supervisionado¹⁵⁵⁸ tem grande potencial, haja vista a aplicabilidade mais intuitiva e facilitada¹⁵⁵⁹ a partir dos conhecimentos e abstrações advindos da figura do assessor-desenvolvedor e dos operadores do Direito.

¹⁵⁵³ HARTMANN PEIXOTO, Fabiano; SILVA, Roberta Zumblick Martins da. Op. cit., p. 91.

¹⁵⁵⁴ “A kind of ML that has been applied to predict legal outcomes is supervised ML. Since it involves inferring a classification model (or function) from labeled training data, the ML is referred to as supervised”. (ASHLEY, Kevin D. **Artificial Intelligence and Legal Analytics**, p. 109).

¹⁵⁵⁵ “Técnicamente, outra condição deve ser atendida para a mineração supervisionada de dados: deve haver dados sobre o alvo. Não é suficiente que as informações alvo existam no princípio; elas também devem existir nos dados. Por exemplo, pode ser útil saber se determinado cliente permanecerá por, pelo menos, seis meses, porém, se no histórico de dados essa retenção de informações estiver ausente ou incompleta (se, por exemplo, os dados são retidos apenas por dois meses) os valores alvo não podem ser fornecidos. A aquisição de dados sobre o alvo costuma ser o segredo do investimento em data science. O valor da variável alvo para um indivíduo é, muitas vezes, chamado de *rótulo* individual, enfatizando que, muitas vezes (não sempre), deve-se incorrer uma despesa para rotular ativamente os dados” (PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 24).

¹⁵⁵⁶ São diversos os algoritmos, sendo que neste momento não é necessário grande detalhamento para além do conceito de aprendizado supervisionado.

¹⁵⁵⁷ LAGE, Fernanda de Carvalho. Op. cit., p. 77.

¹⁵⁵⁸ “The supervised learning algorithm analyzes the training data and infers a function or model, which ideally can be used to classify new, unseen instances. The model may, for example, be embodied in statistically computed feature weights or in a set of rules mechanically induced from the training data. Given a new instance, the model is applied to predict an outcome”. (ASHLEY, Kevin D. **Artificial Intelligence and Legal Analytics**, p. 109-110).

¹⁵⁵⁹ “Tarefas supervisionadas exigem técnicas diferentes das não supervisionadas, e os resultados costumam ser mais úteis. Uma técnica supervisionada recebe um objetivo específico para o agrupamento – prever o alvo. O agrupamento, uma tarefa não supervisionada, produz grupos baseados

A respeito da dinâmica de atuação do aprendizado supervisionado, veja-se o seguinte gráfico:

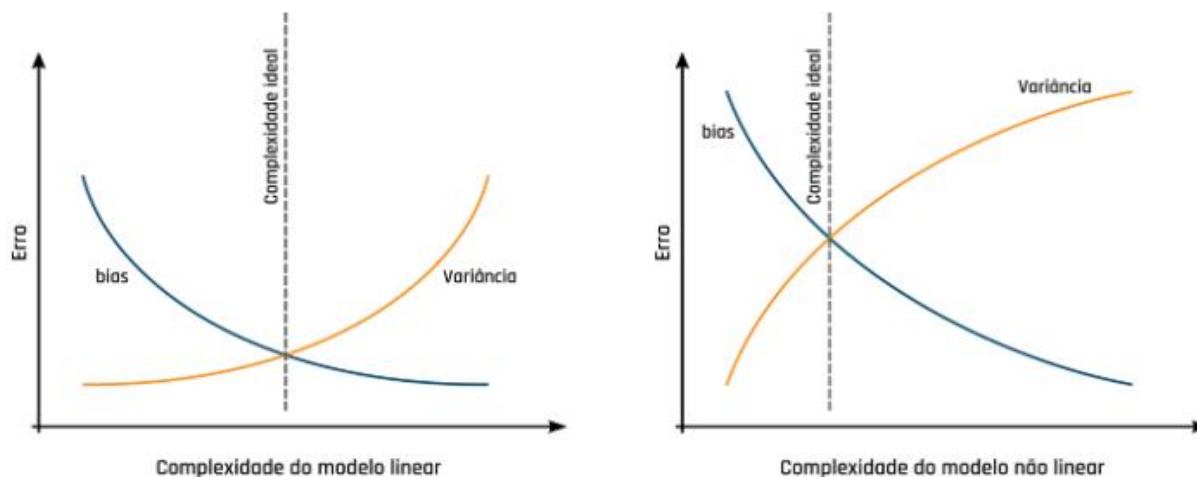


Fig. 6 – Trade-off no aprendizado supervisionado.¹⁵⁶⁰

Já no **aprendizado não supervisionado** os algoritmos exploram os dados de maneira a buscar por padrões, correlações e relacionamentos, sem a necessidade de programação humana voltada à pré-análise dos dados.¹⁵⁶¹ Desse modo, no aprendizado não supervisionado não são pré-estabelecidas variáveis alvo, ou seja, os *outputs* são resultado das inferências automaticamente encontradas pelos algoritmos através das entradas de dados fornecidas (*inputs*). Por exemplo, a busca de correlações que naturalmente acontecem entre variáveis¹⁵⁶² (a partir de dados dos processos, das decisões ou das partes, por exemplo) é uma forma de utilização de técnicas de aprendizado não supervisionado. Todavia, deve ser dedicado cuidado especial no caso de aprendizado não supervisionado, a fim de que inferências

em similaridades, mas não há garantia de que essas semelhanças sejam significativas ou úteis para qualquer propósito em particular” (PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 24).

¹⁵⁶⁰ ALMEIDA, Adriano; CARVALHO, Felipe; MENINO, Felipe; NETO, Helvécio. **Introdução ao Machine Learning**. [2017]. Disponível em: dataat.github.io/introducao-ao-machine-learning/introducao.html. Acesso em 31 jan. 2023.

¹⁵⁶¹ PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 24.

¹⁵⁶² “Considere duas perguntas semelhantes que podemos fazer sobre uma população de clientes. A primeira é: ‘Nossos clientes naturalmente se encaixam em grupos diferente?’ Aqui, nenhuma proposta, ou *alvo*, em particular foi especificada para o agrupamento. Quando não existe alvo, o problema de mineração de dados é chamado de *não supervisionado*. Compare isso com uma pergunta ligeiramente diferente: ‘Podemos encontrar grupos de clientes que tenham probabilidades particularmente elevadas de cancelar seus serviços logo após o vencimento dos seus contratos?’ Aqui há um alvo específico definido: será que um cliente abandonará o serviço quando seu contrato vencer? Neste caso, a segmentação está sendo feita por um motivo específico: tomar medidas com base na probabilidade de rotatividade. Isso é chamado de problema supervisionado de mineração de dados” (Ibid., p. 24).

baseadas em vieses raciais, sociais ou de outros elementos de discriminação arbitrária não sejam replicadas da base de dados (a partir de correlações encontradas) para o modelo de IA judicial.

O aprendizado não supervisionado demonstra sua aplicação pelo fato que tem o potencial de encontrar correlações que passavam despercebidas (ou, até mesmo, que alguém preferisse esconder), o que tem grande valia quando se pensa em sistemas de IA judicial voltados à transparência e à explicabilidade lastreadas no duplo canal de correção.

Abaixo, segue uma representação gráfica do aprendizado não supervisionado:

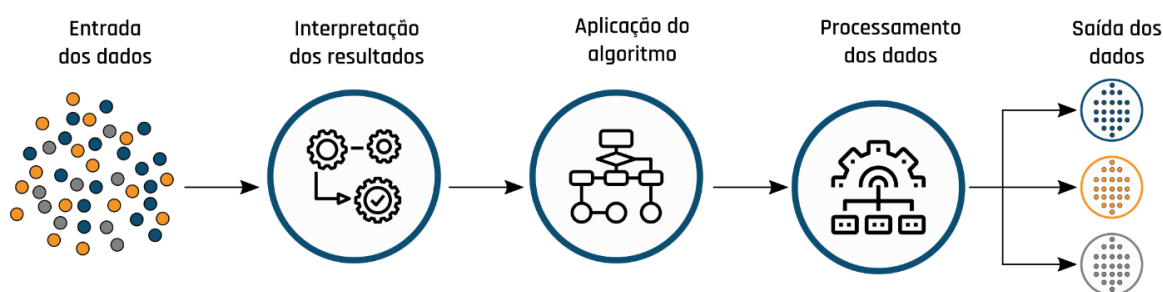


Fig. 7 – Aprendizado não supervisionado.¹⁵⁶³

O último dos três principais paradigmas aqui enfocados é **aprendizagem por reforço**¹⁵⁶⁴, a qual se volta abordagem voltada à interação¹⁵⁶⁵ para que o aprendizado ocorra. Desse modo, no aprendizado por reforço, o sistema busca extrapolar suas reações a fim de que tenha a capacidade de agir corretamente em situações que não estavam na sua base de dados (*dataset*) de treinamento.¹⁵⁶⁶ Nesse modelo, o “aprendizado por reforço busca maximizar um sinal de bom desempenho, e não encontrar uma estrutura oculta”¹⁵⁶⁷, ainda que essa última função possa ser explorada

¹⁵⁶³ ALMEIDA, Adriano; CARVALHO, Felipe; MENINO, Felipe; NETO, Helvécio. Op. cit.

¹⁵⁶⁴ “Considere, por exemplo, o problema de aprender a jogar a xadrez. Um agente de aprendizagem supervisionada precisa ser informado da jogada correta para cada posição que encontra, mas tal realimentação raramente está disponível. Na ausência de realimentação de um professor, um agente pode aprender um modelo de transição para seus próprios movimentos e talvez possa aprender a prever as jogadas do adversário, mas *sem alguma realimentação sobre o que é bom e o que é ruim, o agente não terá nenhuma base para decidir que movimento executar*. O agente precisa saber que algo de bom aconteceu quando (acidentalmente) dá o xeque-mate no oponente ou vice-versa, se for um jogo suicida. Essa espécie de realimentação é chamada de recompensa ou reforço” (RUSSELL, Stuart; NORVIG, Peter. Op. cit., p. 723).

¹⁵⁶⁵ HARTMANN PEIXOTO, Fabiano; SILVA, Roberta Zumblick Martins da. Op. cit., p. 95.

¹⁵⁶⁶ Ibid., p. 96.

¹⁵⁶⁷ Id.

de maneira secundária neste paradigma de *machine learning*. Veja-se o fluxo básico da aprendizagem por reforço:

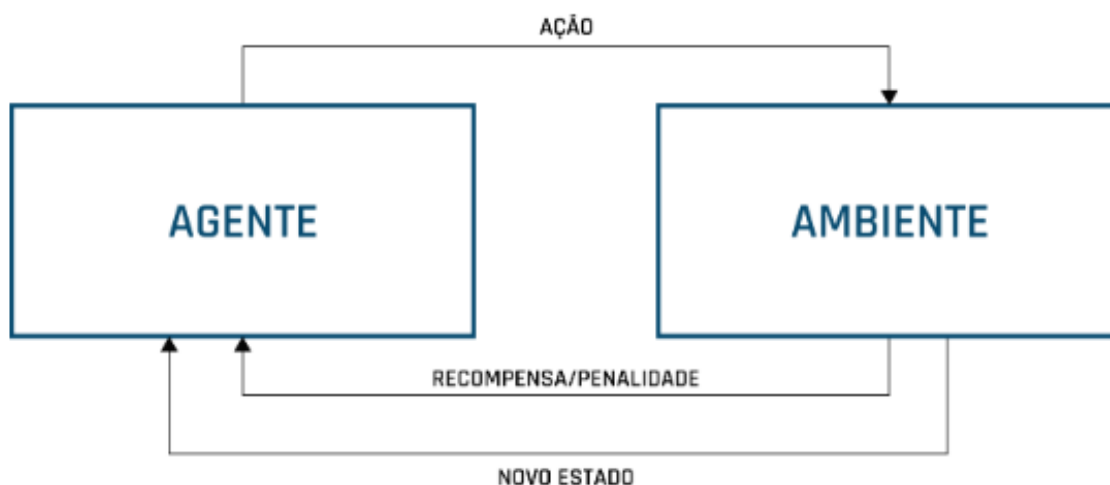


Fig. 8 – Fluxo básico da aprendizagem por reforço.¹⁵⁶⁸

A partir desta breve definição dos paradigmas de aprendizado de máquina (*machine learning*), destaca-se que, geralmente: a) modelos de IA voltados para classificação, regressão e modelagem causal são desenvolvidos através de aprendizagem supervisionada; b) os modelos de Inteligência Artificial destinados para agrupamento, agrupamento por coocorrência e perfilamento utilizam aprendizado não supervisionado; c) já modelos baseados em combinação por similaridade, previsão de vínculo e redução de dados podem decorrer do uso de qualquer dos dois paradigmas referidos.¹⁵⁶⁹

5.6 A Pretensão de Correção no Ciclo da Ciência de Dados Jurídica: O Desenvolvimento de Modelos de *Machine Learning* para a IA Judicial

Neste subcapítulo, será realizada a abordagem do ciclo de desenvolvimento de sistemas de Inteligência Artificial judicial baseados em *machine learning* e sua relação com o duplo canal de correção. Dessa maneira, conforme já destacado no início deste Capítulo 5, a leitura que esta obra faz do tema tem o papel de compatibilizar a Ciência da Computação e a Ciência de Dados à Teoria do Direito, motivo pelo qual a visão

¹⁵⁶⁸ ALMEIDA, Adriano; CARVALHO, Felipe; MENINO, Felipe; NETO, Helvécio. Op. cit.

¹⁵⁶⁹ PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 25.

panorâmica dos temas tipicamente relacionados às áreas da Tecnologia da Informação se deve à necessária interdisciplinaridade da pesquisa. Nesse caminhar, em alguns momentos o texto se baseia em simplificações e concatenações de ideias que podem soar estranhas ou fora de local para um leitor não jurídico, mas que têm seu papel estruturante para a compreensão do problema e da obra a partir do Direito. Feito o aviso necessário, passa-se ao tema do tópico.

Há uma inadequada impressão em parte da comunidade jurídica no sentido de que sistemas de IA judicial poderiam ser desenvolvidos nos moldes “*plug and play*”¹⁵⁷⁰ ou algo como: “preparar, apontar, fogo”!

No entanto, a partir da Ciência de Dados e dos estudos da aplicação de Inteligência Artificial no Direito, a referida impressão deve ser afastada a partir do paradigma *Data Science*¹⁵⁷¹, uma vez que cada base de dados explorada para a construção de modelos de aprendizado de máquina tem características que lhes são peculiares. Assim, no desenvolvimento dos sistemas de IA, deve-se “conseguir identificar falhas óbvias, hipóteses fantasiosas e partes faltando” no *dataset*.¹⁵⁷² Nesse caminho, técnica e experiência caminham juntas na busca de se desenvolver sistemas de IA judicial que tenham desempenho adequado e esperado.

O trilhar desse caminho passa pelas fases de:

- a) compreensão do problema em que será aplicada a solução de IA judicial;
- b) preparação dos dados (o que inclui o pré-processamento destes);
- c) modelagem do subsistema de IA judicial;
- d) avaliação do modelo.

Essa abordagem funciona através do *framework* cíclico, conhecido como Processo Padrão de Indústria Cruzada para Exploração de Dados (*CRISP-DM*)¹⁵⁷³, o qual pode ser ilustrado através da figura abaixo:

¹⁵⁷⁰ Numa abordagem empírica, pode-se rememorar os dias em que o *Microsoft Windows 95* iniciou este paradigma na computação pessoal, contudo o insucesso e a frustração dos usuários que tentavam instalar periféricos (monitores, mouses, *joysticks*, etc.) naquela época ainda está na memória (os problemas de compatibilidade de *hardware* ocorriam por vários motivos e não cabem nesta nota de rodapé).

¹⁵⁷¹ [Pode-se definir que] “*data science* é um conjunto de princípios fundamentais que norteiam a extração de conhecimento a partir de dados. *Data mining* é a extração de conhecimento a partir deles, por meio de tecnologias que incorporam esses princípios” (PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 2).

¹⁵⁷² Id.

¹⁵⁷³ Ibid., p. 27.

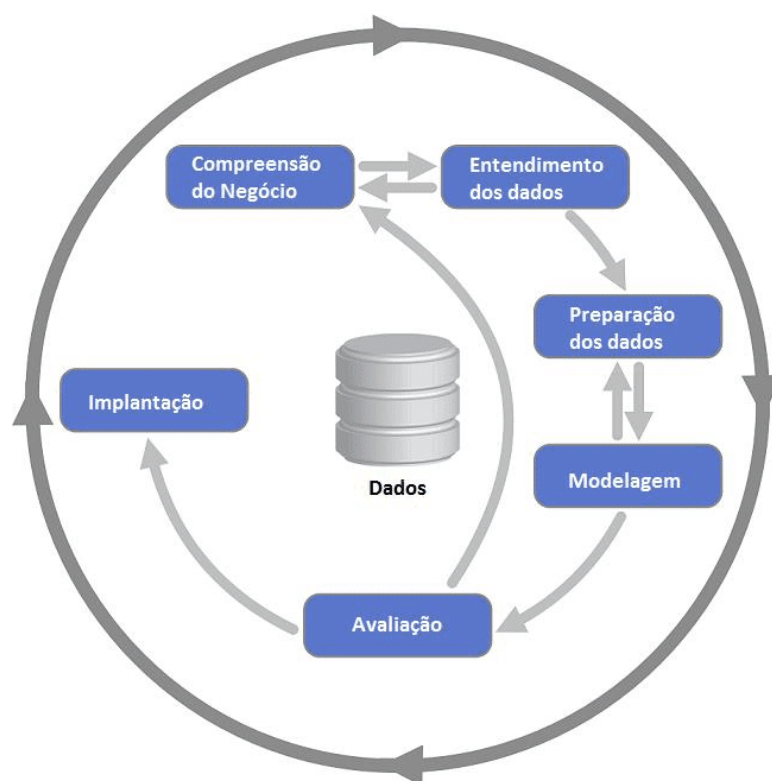


Fig. 9 – Processo Padrão de Indústria Cruzada para Exploração de Dados (CRISP-DM).¹⁵⁷⁴

A partir dessa visualização, percebe-se que um “princípio importante de *data science* é que a mineração de dados é um processo com estágios muito bem definidos”¹⁵⁷⁵, o que passa por um conjunto de tarefas bem delineadas pela decomposição do objetivo global de encontrar padrões no conjunto de dados.¹⁵⁷⁶

Dessa maneira, percebe-se que **cada problema de tomada de decisão pela IA judicial orientada pelos dados tem uma abordagem exclusiva**, a qual tem suas próprias potencialidades e limitações técnicas, ainda que existam “conjuntos de tarefas comuns que permeiam os problemas”¹⁵⁷⁷ a serem explorados. Por exemplo, classificar petições a partir de similaridade em seus argumentos e classificar processos judiciais a partir das características de seus trâmites são tarefas diferentes e que podem ter objetivos bastantes diversos, mas existem tarefas comuns de Ciência de Dados que serão aplicadas a ambos os problemas.

¹⁵⁷⁴ SHEARER, Colin. The CRISP-DM Model: The New Blueprint for Data Mining. **Journal of Data Warehousing**, v. 5, n. 4, 2000.

¹⁵⁷⁵ PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 19.

¹⁵⁷⁶ GRUS, Joel. Op. cit., p. 2.

¹⁵⁷⁷ PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 19.

Nessa abordagem de construção de Inteligência Artificial judicial a partir de dados, o cientista de dados jurídicos¹⁵⁷⁸ (assessor-desenvolvedor¹⁵⁷⁹) deve decompor o problema jurídico através dos dados obtidos, de maneira que possa aplicar as ferramentas de Ciência de Dados em cada uma das tarefas. Essa abordagem de aplicação da Ciência de Dados para o desenvolvimento de IA judicial permite que os demais profissionais do Direito (servidores do Judiciário e magistrados¹⁵⁸⁰) “concentrem sua atenção em partes mais interessantes do processo que requerem envolvimento humano”¹⁵⁸¹, especialmente quando forem necessários o pensamento humano¹⁵⁸² e a sua criatividade inerentemente reflexiva.

Consideradas as peculiaridades de cada fase do ciclo de desenvolvimento de modelos de Inteligência Artificial aplicáveis ao Direito, observa-se que a pretensão de correção¹⁵⁸³ se manifesta em todas as fases, seja em sua vertente normativa (correta interpretação do sistema normativo) ou através da faceta da pretensão de correção moral (abstração de padrões éticos representados nos dados ou ainda a verificação de enviesamento por meio dos dados), o que passa a influenciar o duplo canal de correção inerente ao juiz ciborgue, pois no “âmbito de abertura do direito, tudo aquilo em que se apoia e/ou deve apoiar-se quem aplica o direito para satisfazer a pretensão à correção integra o direito”.¹⁵⁸⁴ Em síntese, **pensar em Ciência de Dados e Inteligência Artificial para o auxílio à decisão judicial é também pensar o Direito.**

¹⁵⁷⁸ “Na verdade, alguns cientistas de dados são – para todos os propósitos práticos – estatísticos, enquanto outros são quase indistinguíveis dos engenheiros de software. Alguns são experts em aprendizado de máquina, enquanto outros não conseguiram aprender muita coisa sobre o assunto. Alguns são PhDs com um impressionante registro de publicações, enquanto outros nunca leram um trabalho acadêmico (apesar de ser uma vergonha). Resumindo, basicamente não importa como você define data science, pois você encontrará praticantes para quem a definição está total ou absolutamente errada” (GRUS, Joel. Op. cit., p. 1-2).

¹⁵⁷⁹ A atuação como cientista de dados é uma faceta das potencialidades do assessor-desenvolvedor.

¹⁵⁸⁰ O destaque para os servidores do Judiciário e para os magistrados ocorre devido ao fato que se está tratando do desenvolvimento de sistemas de IA judicial. Todavia, isso não inibe a colaboração de advogados, de integrantes do Ministério Público e da academia, na forma delineada na Resolução 332/2020 do Conselho Nacional de Justiça.

¹⁵⁸¹ PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 20.

¹⁵⁸² KAHNEMAN, Daniel. Op. cit., posição 281.

¹⁵⁸³ “O direito é um sistema normativo que (1) formula uma pretensão de correção, (2) consiste na totalidade das normas que integram uma constituição socialmente eficaz em termos globais e que não são extremamente injustas, bem como na totalidade das normas estabelecidas em conformidade com essa constituição e que apresentam um mínimo de eficácia social ou de possibilidade de eficácia e não são extremamente injustas, e (3) ao qual pertencem os princípios e outros argumentos normativos, nos quais se apoia e/ou deve se apoiar o procedimento de aplicação do direito para satisfazer a pretensão à correção” (ALEXY, Robert. **Conceito e Validade do Direito**, p. 151).

¹⁵⁸⁴ Ibid., p. 155.

Colocadas estas premissas, destaca-se que a fase de **compreensão do problema jurídico**¹⁵⁸⁵ demanda um aspecto de arte da Ciência de Dados jurídica, momento em que a criatividade¹⁵⁸⁶ do assessor-desenvolvedor¹⁵⁸⁷ e do juiz tem grande importância, pois, para que alcance os objetivos, o problema jurídico deve ser elaborado “como um ou mais problemas de data science”.¹⁵⁸⁸ Neste aspecto, o conhecimento de alto nível¹⁵⁸⁹ sobre o Direito pode nortear a equipe de desenvolvimento a pensar em soluções criativas a partir da Ciência de Dados. Com isso, passa-se a perceber que a fronteira entre conhecer parte do Direito e aplicar Ciência de Dados jurídica começa a ficar bastante estreita, o que pode assustar ou causar estranhamento ao profissional do Direito (e ao jurista¹⁵⁹⁰) que recusar o aprofundamento técnico desse entrelace interdisciplinar.

Nesse caminho, a compreensão do problema jurídico a ser resolvido (que deve ser cuidadosamente pensado pela equipe¹⁵⁹¹) pode levar a outros subproblemas que necessitem do desenvolvimento de modelos de IA judicial voltados para classificação, regressão ou outro paradigma de Ciência de Dados.

Na sequência, passa-se à fase de **compreensão dos dados**, o que depende da prévia existência¹⁵⁹² de um conjunto de dados ou de sua coleta¹⁵⁹³ a partir da definição do problema jurídico, o que necessitará da estimativa de custo, financeiro e humano, para se verificar a viabilidade de cada projeto.¹⁵⁹⁴

¹⁵⁸⁵ Aqui há opção em substituir o termo “compreensão do negócio” (típico da abordagem *Data Science*) para “compreensão do problema jurídico”, a fim de manter a ambientação do leitor nas questões do Direito.

¹⁵⁸⁶ “Inicialmente, é vital compreender o problema a ser resolvido. Isso pode parecer óbvio, mas projetos de negócios raramente vêm pré-moldados como problemas claros e inequívocos de mineração de dados. Muitas vezes, reformular o problema e projetar uma solução é um processo repetitivo de descoberta”. (PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 27).

¹⁵⁸⁷ [É] “preciso encontrar pessoas – provavelmente de sua organização – que tenham o conhecimento de domínio adequado para o projeto de IA. Elas precisarão refletir sobre os fluxos de trabalho, modelos e dados de treinamento – com a compreensão específica dos requisitos do setor” (TAULLI, Tom. Op. cit., p. 52).

¹⁵⁸⁸ PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 28.

¹⁵⁸⁹ Id.

¹⁵⁹⁰ E aqui, neste momento de bastante estranhamento com o novo, está um dos papéis pedagógicos do presente estudo.

¹⁵⁹¹ PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 28.

¹⁵⁹² “Os dados históricos, muitas vezes, são recolhidos para fins não relacionados com o problema de negócio atual ou para nenhum propósito explícito” (Id.).

¹⁵⁹³ “Alguns dados estarão disponíveis praticamente de graça, enquanto outros exigirão um esforço para serem obtidos” (Id.).

¹⁵⁹⁴ “Uma parte crítica da fase de compreensão dos dados é estimar os custos e benefícios de cada fonte e decidir se ainda mais investimentos são necessários. Mesmo depois de todos os conjuntos de dados serem adquiridos, **conferi-los pode exigir esforço adicional**” (Ibid., p. 28-29).

Para que haja a compreensão dos dados jurídicos, é necessário revelar a estrutura do problema jurídico através do conjunto de dados obtido (e as técnicas de **visualização** de dados¹⁵⁹⁵ são auxiliares importantes neste processo), a fim de inferir quais tecnologias de mineração de dados podem ser aplicadas em cada caso. Nessa fase, deve ser destacado que um problema jurídico representado através de seus dados pode exigir diversas tarefas de mineração de dados, sendo, muitas vezes, necessário combinar tipos diferentes de soluções¹⁵⁹⁶ para se alcançar um desempenho satisfatório.

Por sua vez, a fase de **preparação dos dados** consiste no necessário pré-processamento a partir de inferências da equipe de desenvolvimento (assessor-desenvolvedor e magistrado¹⁵⁹⁷) na conversão de dados¹⁵⁹⁸, na eliminação de séries de dados faltantes e em outras tarefas de pré-processamento, de maneira que as tecnologias de Ciência de Dados possam atuar com eficiência para alcançar o melhor desempenho possível no treinamento.

Em relação aos textos existentes nas bases de dados, são exemplos de tarefas de preparação dos dados a **a) normalização**¹⁵⁹⁹, que consiste na conversão das palavras para minúsculas e redução das palavras aos seus prefixos básicos); a **b) tokenização**¹⁶⁰⁰, que se dá através da retirada da pontuação e dos caracteres especiais (para que cada palavra delimitada corresponda um *token*), da junção de palavras (n-gramas), para sejam lidas como se fossem uma palavra só, e da

¹⁵⁹⁵ GRUS, Joel. Op. cit., p. 37-47.

¹⁵⁹⁶ PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 29.

¹⁵⁹⁷ As figuras do assessor-desenvolvedor e do magistrado compõe uma equipe mínima, embora a equipe possa ser maior, a depender a da extensão do problema e dos recursos humanos e financeiros disponíveis.

¹⁵⁹⁸ PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 29-30.

¹⁵⁹⁹ “Normalization involves converting words to lower case and stemming them to their uninflected roots in order to eliminate superficial variations (Turney and Pantel, 2010, p.154). For example, ‘Phelps’ is converted to ‘phelps.’ ‘Swimming’ and ‘swam’ are converted to ‘swim’” (ASHLEY, Kevin D. **Artificial Intelligence and Legal Analytics**, p. 236).

¹⁶⁰⁰ “Tokenizing texts involves transforming words that are hyphenated, have apostrophes, or have other punctuation, into standard forms. Tokenizing also frequently involves eliminating stop words, “high-frequency words with relatively low information content, such as function words (e.g., of, the, and) and pronouns (e.g., them, who, that)” (Turney and Pantel, 2010, p. 154).

Tokenization can present some subtleties. For instance, “Michael” (with a comma) is commonly transformed to “michael.” “Michael’s” with an apostrophe, however, may be a possessive form or a contraction. Different text processing systems employ different rules for tokenizing words like this, which may or may not preserve the semantic distinction. The choice of stop words can also have an effect on meaning. In some contexts, for example web search, frequently appearing short words like “and,” “any,” “not,” and “or” may be treated as stop words, but in a legal context they often convey significant information content. In practice, text processing systems implement default rules for tokenization, which may need to be examined in light of the legal application” (Id.).

eliminação de *stopwords* (palavras com alta taxa de repetição que possuem pouco conteúdo informativo¹⁶⁰¹); e a **c) anotação**¹⁶⁰², que ocorre mediante a inclusão de informações para fins eliminar ou reduzir a ambiguidade de palavras semelhantes no texto, como anotações de sentido ou da função gramatical, por exemplo. A partir desses exemplos, pode-se perceber que cada tarefa de preparação de dados é única e tem suas próprias peculiaridades voltadas à transformação do “*corpus textual* (um conjunto de documentos, tais como acervo de processos de um tribunal), em um espaço vetorial multidimensional”¹⁶⁰³ no qual os algoritmos possam atuar.

Também se faz necessária a **engenharia de variáveis**, pois, por exemplo, pode ser necessário transformar dados representados por texto (*strings*) para dados numéricos (inteiros ou flutuantes¹⁶⁰⁴), a fim de que seja possível sua utilização por algoritmos de aprendizado de máquina, como a máquina de vetores de suporte (SVM¹⁶⁰⁵). Além disso, podem ser necessários outros procedimentos preparatórios para possibilitar a aplicação de tecnologias de *machine learning*, como: a **categorização de variáveis contínuas**¹⁶⁰⁶; o estabelecimento de **escalas para variáveis contínuas** (através de técnicas de **padronização**¹⁶⁰⁷ ou de **normalização**¹⁶⁰⁸); e a transformação de variáveis categóricas¹⁶⁰⁹ em **variáveis**

¹⁶⁰¹ Os artigos “a”, “o”, “os” e “as”, por exemplo.

¹⁶⁰² “Annotation involves adding information that may help to disambiguate similar words in a text. Useful annotations include the sense in which an ambiguous word is used in the text (sense tagging). Other annotations are words’ parts of speech in the text, that is, nouns, verbs, adjectives, or adverbs (POS tagging). One can also annotate the words’ grammatical roles in the parsed sentences’ structures, such as direct object or object of a prepositional phrase (see Turney and Pantel, 2010, p. 155)” (ASHLEY, Kevin D. **Artificial Intelligence and Legal Analytics**, p. 237).

¹⁶⁰³ BOEING, Daniel Henrique Arruda; DA ROSA, Alexandre Moraes. Op. cit., p. 95.

¹⁶⁰⁴ Em Python representados por “*int*” (1, 2 ou 3, por exemplo) e “*float*” (2.3, 4.1 ou 11.5, por exemplo), respectivamente.

¹⁶⁰⁵ Sigla para *Support Vector Machine*.

¹⁶⁰⁶ SILVA, Diego Mattozo Bernardes da. **Métodos de categorização de variáveis preditoras em modelos de regressão para variáveis binárias**. 2017. Dissertação (Mestrado em Estatística). Universidade Federal de São Carlos (UFScar), São Carlos, 2017.

¹⁶⁰⁷ “Já o bloco Análise exploratória de dados tem o objetivo de entender e explorar os dados disponíveis, o que inclui a geração de estatísticas descritivas e visualizações gráficas dos dados; e o bloco Pré-processamento inclui tarefas de preparação dos dados como limpeza de dados (como remoção de duplicatas, tratamento de valores faltantes), seleção de atributos (removendo, por exemplo, atributos redundantes ou pouco relevantes) e transformação de dados (por exemplo, normalização e padronização de dados)” (ESCOVEDO, Tatiana; KOSHIYAMA, Adriano. Op. cit., posição 3500-3506).

¹⁶⁰⁸ “A normalização é um recurso útil quando as quantidades são menos significativas que os valores relativos. Por exemplo, podemos estar menos interessados na idade absoluta do cliente do que o quanto mais velho ou mais novo ele é em relação a um cliente típico. Neste caso, a idade dos clientes pode ser normalizada pela média das idades: quanto maior o valor da variável normalizada, mais velho é o cliente” (Ibid., posição 1331-1337).

¹⁶⁰⁹ Variável que pode assumir um número limitado de valores possíveis. Por exemplo, o estado civil de uma pessoa pode ser “solteira”, “casada”, “em união estável”, “divorciada” ou “viúva”.

dummy (variáveis numéricas). Ademais, destaca-se que a rápida referência às técnicas de preparação de dados serve como aparato inicial ao profissional do Direito, o que, guardadas as devidas proporções, também é função teórica desta obra¹⁶¹⁰, pois isso se faz necessário para a compreensão de todo o contexto da tese desenvolvida (duplo canal correção inerente à atuação do modelo de juiz ciborgue).

Além disso, a eliminação de **outliers**¹⁶¹¹ é parte importante da preparação de dados, pois os *outliers* podem causar falhas na predição dos modelos de IA judicial, uma vez que influenciam o processo de aprendizado de máquina de maneira a causar queda de desempenho ou sobreajuste (*overfitting*), o que leva a modelos da IA baseados em aprendizado de máquina que têm baixa capacidade de generalização nos dados da vida real. Um *outlier* pode decorrer de erro de medição ou de entrada (*input*), de corrupção dos dados ou de verdadeira observação atípica.

A fase de preparação dos dados envolve bastante trabalho humano especializado em Direito¹⁶¹², e que tenha pretensão de correção, pois numa fase inicial de qualquer projeto é necessário tempo, criatividade humana e conhecimento jurídico para definir as variáveis que serão utilizadas nas fases seguintes, pois comumente “a qualidade da solução de mineração de dados baseia-se em quão bem os analistas estruturam os problemas e criam variáveis”.¹⁶¹³

Dessa forma, para além das pesquisas estatísticas realizadas pelos cientistas de dados jurídicos, a verificação e o saneamento de *outliers* devem ser realizadas pelo magistrado e pelo assessor-desenvolvedor, a fim de analisar a sua pertinência em relação ao conjunto de dados. Destaca-se que a manutenção de um estado de pretensão de correção só é possível com a participação legítima do julgador na busca do adequado desempenho do sistema de IA judicial, razão pela qual, no exemplo do tratamento de *outliers*, não seria legítima a exclusão de dados (a sua consideração como ineficientes para a modelagem da IA) à margem da efetiva participação do

¹⁶¹⁰ Cabe dizer, mais uma vez, que não há como se pensar no entrelaçamento entre decisão judicial e inteligência artificial sem a referência aos conceitos e às ferramentas básicas da Ciência da Computação e da Ciência de Dados. A parte técnica pode causar “dores de cabeça” aos profissionais do Direito, mas varrer o problema para “debaixo do tapete” e soltar “fogos de artifício” com uma escrita puramente jurídica não daria conta da questão. Este aviso é reiterado diversas vezes durante texto do presente estudo, mas se mostra sempre necessário.

¹⁶¹¹ “Dados discrepantes (*outliers*): estão bem além do escopo da maior parte dos demais dados; o que pode indicar não serão úteis. Há situações, no entanto, em que o inverso é verdadeiro; o que serviria para detecção de fraudes” (TAULLI, Tom. Op. cit., p. 54).

¹⁶¹² Id.

¹⁶¹³ PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 30.

jugador num contexto de cooperação humano-máquina no modelo de juiz ciborgue aqui desenvolvido.

Para além dos aspectos mais gerais, um cuidado adicional que deve ser tomado no que se refere à rotulagem de dados voltados ao aprendizado de máquina para sistemas de aferição de viés decisório do julgador, pois, nesse caso, para que haja a investigação computacional de vieses do juiz na aplicação do Direito, o próprio julgador que será analisado pelo sistema não é a única e a melhor fonte para rotular os dados, pois pode, em decorrência dos próprios vieses psicológicos (sem intenção), atuar de maneira a diminuir a representatividade da base de dados em relação à realidade, o que diminui o desempenho dos sistemas de IA judicial ou os inviabiliza, a depender de cada caso. Nessa hipótese, deve sempre existir equipe interdisciplinar que, além de ser integrada pelo magistrado, conte com desenvolvedores e outros profissionais do Direito que não serão afetados diretamente pelo sistema de IA judicial.

Por outro lado, quando o sistema de aferição de viés decisório for desenvolvido no âmbito recursal com o objetivo de aferir a coerência dos juízes cujas decisões são recorridas, é recomendável que os magistrados que tiverem suas decisões submetidas às técnicas de perfilamento tenham conhecimento e acesso à fase de preparação dos dados, a fim de que possam contribuir para o desenvolvimento do sistema de IA judicial através de questionamentos e sugestões de aprimoramento.

Mais uma vez, tais premissas de desenvolvimento da IA judicial são inerentes ao duplo canal de correção pela via comunicacional humano-máquina, no modelo de entrelace entre a pretensão de correção (Robert Alexy¹⁶¹⁴) e o canal de correção (Shannon¹⁶¹⁵) que está no cerne da presente obra (vide subcapítulo 1.6).

A fase seguinte é **modelagem**, que consiste na construção de “algum tipo de modelo ou padrão que captura regularidade nos dados”¹⁶¹⁶, ou, em outras palavras, é “a especificação de uma relação matemática (ou probabilística) existente entre variáveis diferentes”.¹⁶¹⁷

Na modelagem, há a utilização de uma ou mais tecnologias de Ciência de Dados para a definição de um modelo IA, o que ocorre a partir da **análise de uma ou**

¹⁶¹⁴ “Sistemas normativos que não formulam explícita ou implicitamente uma pretensão de correção não são sistemas jurídicos” (ALEXY, Robert. **Conceito e Validade do Direito**, p. 152).

¹⁶¹⁵ SHANNON, Claude E. *A Mathematical Theory of Communication*.

¹⁶¹⁶ PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 31.

¹⁶¹⁷ GRUS, Joel. Op. cit., p. 141.

mais variáveis de predição, de modo a estimar a probabilidade de ocorrência (classificação) ou o valor quantitativo (regressão) de **uma determinada variável alvo**.

Por exemplo, para auxiliar no julgamento dos requisitos para a concessão dos benefícios de gratuidade de justiça¹⁶¹⁸ (variável alvo: justiça gratuita), podem ser analisadas variáveis preditivas (salário líquido, tipo de atividade econômica ou laborativa, nível de escolaridade, número de filhos, despesas para a manutenção da família, etc.) a partir do conjunto de dados preparado. Levando-se em conta esse exemplo, pode existir a construção de um modelo de aprendizado de máquina que responda a probabilidade da parte ter direito à justiça gratuita (variável alvo), o que pode ser utilizado pelo julgador no auxílio à decisão sobre o tema.

Sendo assim, percebe-se que a **variável alvo** é a informação que se busca prever ou investigar mediante a classificação (probabilidade do acontecimento do evento) ou a regressão (probabilidade da intensidade do evento). Já as **variáveis preditoras** são as que decorrem de dados históricos conhecidos existentes em uma base de dados (*dataset*), as quais serão exploradas para se estabelecer o grau de correlação com a variável alvo e, na sequência, para a modelagem do sistema de IA judicial.

São vários os algoritmos que podem ser utilizados para a construção de modelos de *machine learning*, sendo que alguns dos exemplos mais conhecidos são: regressão linear¹⁶¹⁹, regressão múltipla¹⁶²⁰, regressão logística¹⁶²¹, *Naive Bayes*¹⁶²², árvore de decisões¹⁶²³ (e suas variantes¹⁶²⁴), K-Vizinhos Mais Próximos (*KNN*¹⁶²⁵)¹⁶²⁶

¹⁶¹⁸ Art. 98 e art. 99 do Código de Processo Civil.

¹⁶¹⁹ RUSSELL, Stuart; NORVIG, Peter. Op. cit., p. 626-628.

¹⁶²⁰ Ibid., p. 628-630.

¹⁶²¹ “A maximum entropy classifier in language processing, also known as multinomial logistic regression, assigns a class to a document by learning what features from the input are most useful for discriminating between the different classes. It computes a probability from an exponential function of a weighted set of these observed features (Jurafsky and Martin, 2015). It is based on the principle that when nothing else is known, the probability distribution should be as uniform as possible, that is, it should have maximum entropy” (ASHLEY, Kevin D. **Artificial Intelligence and Legal Analytics**, p. 290).

¹⁶²² TAULLI, Tom. Op. cit., p. 81-83.

¹⁶²³ PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 62-78.

¹⁶²⁴ *Random Forest*, por exemplo: “*RFDT replaces a single decision tree with ensembles or random forests of decision trees in order to achieve greater diversity of sources in prediction. “Ensemble” methods generate a number of diverse trees and average across an entire forest*” (ASHLEY, Kevin D. **Artificial Intelligence and Legal Analytics**, p. 113).

¹⁶²⁵ Sigla para K - *Nearest Neighbors*.

¹⁶²⁶ PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 147-151.

e máquina de vetores de suporte¹⁶²⁷ (SVM).¹⁶²⁸ Esta obra não realizará um aprofundamento nos algoritmos voltados à modelagem de IA, tendo em vista que não é o objetivo primário deste estudo discutir algoritmos específicos, mas sim as implicações da IA para o Direito.¹⁶²⁹

Neste ponto, ainda que se possa imaginar que seria de grande dificuldade programar os algoritmos para o desenvolvimento de modelos IA voltados ao Direito¹⁶³⁰, esse tipo de receio é, em certa medida, exagerado¹⁶³¹, uma vez que em linguagens como Python, por exemplo¹⁶³², existem diversas bibliotecas¹⁶³³ (por

¹⁶²⁷ “An SVM identifies a boundary in a vector space between the positive and negative instances of a category or class that is maximally distant from all of the instances (Noble, 2006). Ideally, all of the candidate cases that are positive instances lie on one side of the boundary and all of the negative instances lie on the other” (ASHLEY, Kevin D.. **Artificial Intelligence and Legal Analytics**, p. 251).

¹⁶²⁸ RUSSELL, Stuart; NORVIG, Peter. Op. cit., p. 648-652.

¹⁶²⁹ Para o aprofundamento na questão, a bibliografia citada no presente estudo é um bom começo ao leitor que se interessar sobre a forma de funcionamento de cada tipo de algoritmo voltado ao aprendizado de máquina.

¹⁶³⁰ E o autor imagina que os profissionais do Direito serão a maioria do público para este texto, embora os leitores da área da Tecnologia da Informação também possam se aventurar pelo Direito a partir do estudo desenvolvido.

¹⁶³¹ “Contudo, não é necessário um exército de engenheiros de alto nível para um projeto de IA. Na verdade, está ficando cada vez mais fácil aplicar *machine learning* e modelos de *deep learning* por conta de sistemas de código aberto como o TensorFlow e plataformas baseadas em nuvem de empresas como Google, Amazon.com e Microsoft. Em outras palavras, serão necessárias apenas algumas pessoas com experiência em ciência de dados” (TAULLI, Tom. Op. cit., p. 51-52).

¹⁶³² E aqui este autor traz exemplo da linguagem Python a partir do seu campo de conhecimento técnico em programação e Ciência e Dados (o olhar de um estudioso do Direito que também passou a estudar Ciência de Computação), embora se possa desenvolver modelos de Inteligência Artificial em outras linguagens de programação. Por exemplo, há a possibilidade de desenvolvimento de modelos de *deep learning* mais eficientes com o uso de linguagem C++, haja vista que normalmente esta tem um melhor desempenho computacional em relação à linguagem Python (ainda que Python seja mais fácil de aprender e começar a usar, um dos motivos de sua popularidade para projetos de Inteligência Artificial). O R é outro exemplo de linguagem tipicamente utilizada para aprendizado de máquina, embora conte com uma comunidade menor de desenvolvedores (mais ligada ao campo da estatística). Logo, tome-se a referência à linguagem Python apenas como um exemplo limitado para facilitar a compreensão do leitor, ainda que este autor recomende fortemente aos futuros assessores-desenvolvedores o estudo de Python e *Data Science*. A grande maioria da comunidade jurídica ainda não compreende as potencialidades benéficas da Inteligência Artificial no Direito (embora intuitivamente já consiga perceber parte dos riscos), motivo pelo qual só uma geração de profissionais com conhecimento híbrido em Direito e Ciência da Computação vai poder dar as respostas que são tão cobradas neste momento da história que marca o ano de 2023.

¹⁶³³ No contexto de linguagens de programação, uma biblioteca é um conjunto de módulos que, por sua vez, são constituídos de diversas sub-rotinas (métodos) pré-programados na linguagem respectiva, os quais podem ser “chamados” para executar operações cuja programação “do zero” seria mais trabalhosa.

exemplo: Pandas¹⁶³⁴, Numpy¹⁶³⁵, Scikit-learn¹⁶³⁶, Datetime¹⁶³⁷ e PyTorch¹⁶³⁸) que fornecem seus módulos (subdivisões da própria biblioteca) e métodos¹⁶³⁹ (característica inerente à Programação Orientada a Objetos¹⁶⁴⁰) para a aplicação dos algoritmos mais conhecidos. Desse modo, a partir de bibliotecas já desenvolvidas são necessários ajustes de parâmetros a serem levados em conta pelos algoritmos (as variáveis preditoras, a variável alvo, o tamanho máximo da árvore de decisão¹⁶⁴¹, etc.), o que demanda mais conhecimento de Ciência de Dados e inferências a partir do Direito e menos programação pura. Um exemplo de biblioteca *low-code* bastante conhecida do Python é a PyCaret, a qual é dividida em dezoito algoritmos de classificação¹⁶⁴² e vinte e quatro algoritmos de regressão¹⁶⁴³, sendo que sua utilização pode ser orientada através da documentação da biblioteca disponível no *website* oficial da comunidade de desenvolvimento.¹⁶⁴⁴

¹⁶³⁴ NUMFOCUS Inc. **Pandas documentation**. [2022]. Disponível em: <https://pandas.pydata.org/docs>. Acesso em 15 nov. 2022.

¹⁶³⁵ NUMFOCUS Inc. **Numpy documentation**. [2022]. Disponível em: <https://numpy.org/doc>. Acesso em 15 nov. 2022.

¹⁶³⁶ SCIKIT-LEARN. **Scikit-learn**: Machine Learning in Python. [2022]. Disponível em: <https://scikit-learn.org>. Acesso em 15 nov. 2022.

¹⁶³⁷ PYTHON Software Foundation. **Datetime**. [2022]. Disponível em: <https://docs.python.org/pt-br/3/library/datetime.html>. Acesso em 15 nov. 2022.

¹⁶³⁸ THE LINUX Foundation. **PyTorch**. [2022]. Disponível em: <https://pytorch.org/docs/stable/index.html>. Acesso em 15 nov. 2022.

¹⁶³⁹ “O outro tipo de referências a atributos de instância é o ‘método’. Um método é uma função que ‘pertence’ a um objeto instância. (Em Python, o termo método não é aplicado exclusivamente a instâncias de classes definidas pelo usuário: outros tipos de objetos também podem ter métodos. Por exemplo, listas possuem os métodos `append`, `insert`, `remove`, `sort`, entre outros. Porém, na discussão a seguir, usaremos o termo método apenas para se referir a métodos de classes definidas pelo usuário. Seremos explícitos ao falar de outros métodos)” (PYTHON Software Foundation. **Classes**. [2022]. Disponível em: <https://docs.python.org/pt-br/3/tutorial/classes.html>. Acesso em 15 nov. 2022).

¹⁶⁴⁰ Em programação o conceito de “método” em linguagem Java equivale ao conceito de “função” em Python. Desse modo: “Os métodos são sub-rotinas que tratam os dados definidos pela classe e, em muitos casos, dão acesso a esses dados. Quase sempre, outras partes do programa interagem com uma classe por seus métodos.

Um método contém uma ou mais instruções. Em um código Java bem escrito, cada método executa apenas uma tarefa” (SCHILDT, Herbert. Op. cit., p. 107).

¹⁶⁴¹ Na hipótese de utilização deste algoritmo de *machine learning*.

¹⁶⁴² *Logistic Regression, K Neighbors Classifier, Naive Bayes, Decision Tree Classifier, SVM - Linear Kernel, SVM - Radial Kernel, Gaussian Process Classifier, MLP Classifier, Ridge Classifier, Random Forest Classifier, Quadratic Discriminant Analysis, Ada Boost Classifier, Gradient Boosting Classifier, Linear Discriminant Analysis, Extra Trees Classifier, Extreme Gradient Boosting, Light Gradient Boosting Machine e Dummy Classifier*.

¹⁶⁴³ *Linear Regression, Lasso Regression, Ridge Regression, Elastic Net, Least Angle Regression, Lasso Least Angle Regression, Orthogonal, Bayesian Ridge, Automatic _bayes.ARDRegression, Passive, Random Sample Consensus, TheilSen Regressor, Huber Regressor, Kernel Ridge, Support Vector Regression, K Neighbors Regressor, Decision Tree Regressor, Random Forest, Extra Trees Regressor, AdaBoost Regressor, Gradient Boosting Regressor, MLP Regressor, Extreme Gradient Boosting, Light Gradient Boosting Machine e Dummy Regressor*.

¹⁶⁴⁴ PYCARET. **PyCaret Official**. [2022]. Disponível em: <https://pycaret.gitbook.io/docs>. Acesso em 15 nov. 2022.

No mais, existem soluções disponíveis para contratação de computação em nuvem voltada à modelagem de Inteligência Artificial, como na *Microsoft Azure*¹⁶⁴⁵, na *AWS*¹⁶⁴⁶, no *Google Cloud*¹⁶⁴⁷ e na *IBM*¹⁶⁴⁸. Por outro lado, deve ser destacado que os riscos dessa opção de desenvolvimento de IA precisam ser bastante ponderados (o respeito ao direito à explicabilidade, por exemplo – vide o subcapítulo 2.6), haja vista que os dados existentes nos *datasets* do Judiciário são sensíveis¹⁶⁴⁹ (dados pessoais e financeiros de partes, por exemplo), o que pode não recomendar o uso destes sistemas terceiros pelo Judiciário (mas que não impossibilita totalmente o seu uso por escritórios de advocacia).

Sendo assim, levando em consideração que existem diversos algoritmos com os quais o desenvolvimento de uma IA judicial pode ser modelado através de aprendizado de máquina, percebe-se que não há “o melhor algoritmo” para aplicação de *machine learning* ao Direito e às questões do Judiciário. O que define isso é o problema jurídico a ser tratado através da IA, numa mistura de técnica, pesquisa exploratória e arte da equipe de desenvolvimento, conforme já descrito anteriormente. Um exemplo disso é que, a depender da solução de aprendizado de máquina a que se chegue, pode existir inclusive uma combinação de modelos¹⁶⁵⁰ (um classificador de análise de precedentes judiciais que utilize uma ponderação entre um modelo de *Random Forest* e um modelo *KNN*, por exemplo) na busca de melhor representar uma solução para a computação jurídica pretendida.

A já comentada divisão do *dataset* em **dados de treino** e **dados de teste** ocorre na fase de modelagem¹⁶⁵¹, de maneira que o modelo realize sua aprendizagem

¹⁶⁴⁵ MICROSOFT. **Azure AI**. [2022]. Disponível em: <https://azure.microsoft.com/pt-pt/solutions/ai>. Acesso em 15 nov. 2022.

¹⁶⁴⁶ AWS. **Adicione facilmente inteligência às suas aplicações**. [2022]. Disponível em: <https://aws.amazon.com/pt/machine-learning/ai-services>. Acesso em 15 nov. 2022.

¹⁶⁴⁷ GOOGLE Cloud. **Produtos de IA e machine learning**. [2022]. Disponível em: <https://cloud.google.com/products/ai>. Acesso em 15 nov. 2022.

¹⁶⁴⁸ IBM. **Soluções de IA**. [2022]. Disponível em: <https://www.ibm.com/br-pt/artificial-intelligence>. Acesso em 15 nov. 2022.

¹⁶⁴⁹ Segundo o art. 5º, II, da Lei Geral de Proteção de Dados: “II – dado pessoal sensível: dado pessoal sobre origem racial ou étnica, convicção religiosa, opinião política, filiação a sindicato ou a organização de caráter religioso, filosófico ou político, dado referente à saúde ou à vida sexual, dado genético ou biométrico, quando vinculado a uma pessoa natural” (BRASIL. **Lei 13.709, de 14 de agosto de 2018**).

¹⁶⁵⁰ “Trata-se, portanto, de uma abordagem que combina diferentes origens e algoritmos teóricos, como mineração de dados e metodologias de computação leve. Portanto, a hibridização das técnicas de IA pode aumentar seu desempenho individual e fazê-los obter muito mais sucesso ao lidar com problemas complexos e em larga escala do que seus componentes individuais” (LAGE, Fernanda de Carvalho. Op. cit., p. 59).

¹⁶⁵¹ Por exemplo: “O fluxo de um problema de Classificação pode ser representado da seguinte forma: a partir de uma base de dados rotulada (para cada exemplo, conhecemos a sua respectiva classe),

nos dados de treino e, posteriormente, possa generalizar respostas para os dados de teste aos quais não teve prévio acesso.

Ainda na fase de modelagem, deve-se atentar para o risco de **sobreaajuste** (*overfitting*)¹⁶⁵², o qual se dá pela ausência de generalização do modelo de Inteligência Artificial baseada em aprendizado de máquina, haja vista que o modelo é incapaz de perceber a variabilidade do contexto de mundo em que atua. Por outro lado, num cenário de *underfitting* o modelo de aprendizado de máquina apresenta baixo desempenho já na fase treinamento, uma vez que não tem aptidão para perceber os relacionamentos existentes entre as variáveis. Em caso de sub-ajuste (*underfitting*)¹⁶⁵³, o modelo de *machine learning* é abandonado sem que haja a necessidade de experimentação na base de dados de teste, haja vista sua inutilidade.

Em outras palavras, em modelos baseados em *machine learning*, o sobreajuste¹⁶⁵⁴ verifica-se quando se tem um alto desempenho no conjunto de dados de treino (tornando o modelo de IA em um especialista enviesado¹⁶⁵⁵ em identificar os padrões dos dados conhecidos), sendo que, por outro lado, o modelo não consegue atuar em dados desconhecidos, vindo a apresentar baixa acurácia. Em suma, quando ocorre *overfitting* o modelo aprende muitíssimo bem os dados de treino (ao ponto de “decorar” as relações existentes no conjunto de treino), todavia apresenta um desempenho bastante ruim quando atua nos dados de teste que lhes desconhecidos, não tendo utilidade frente a novos dados para os quais não foi treinado.

O sobreajuste é um fenômeno típico do **dilema viés x variância**¹⁶⁵⁶, pois alguma variabilidade dos dados de treinamento é necessária, de maneira a

geram-se dois subconjuntos disjuntos: a base de treino (contendo, por exemplo, 70% dos dados originais) e a base de teste (contendo, por exemplo, 30% dos dados originais), de acordo com as estratégias de amostragens de dados discutidas no capítulo 4” (ESCOVEDO, Tatiana; KOSHIYAMA, Adriano. Op. cit., posição 1374).

¹⁶⁵² PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 111-113.

¹⁶⁵³ GRUS, Joel. Op. cit., p. 143.

¹⁶⁵⁴ “Overfitting occurs when an ML model has so many extra terms that it fits the random variations in data rather than real patterns (see NIST/SEMATECH, 2016). In effect, the model “memorizes” the particular training data rather than “learns” a generalization from it. Such a model will fail when making predictions about new data; it has not learned a generalized classifier from the training data that it can apply to new data.

Some ML models are especially prone to overfitting, including decision trees, which almost always overfit the data. Combining the smaller trees of a randomized forest of decision trees protects against overfitting better than a single complex tree” (ASHLEY, Kevin D. **Artificial Intelligence and Legal Analytics**, p. 113).

¹⁶⁵⁵ TALEB, Nassim N. Op. cit., p. 195.

¹⁶⁵⁶ “Outra maneira de pensar sobre o problema de sobreajuste é um compromisso entre polarização e variância” (GRUS, Joel. Op. cit., p. 147).

representar o mundo real, sob risco de ocorrer o *overfitting*. Assim, quando se pensa em IA, o sobreajuste pode se manifestar, por exemplo: pelo excesso de variáveis preditoras; pelo excesso de vizinhos próximos (modelo *KNN*); ou pelo exacerbado número de nós em uma árvore de decisão. A lista está muito longe de ser exaustiva, apenas a partir de cada projeto de IA (cada problema jurídico é único em termos computacionais) se pode inferir qual é a configuração que apresenta o melhor desempenho, pois não existe uma resposta pronta padronizada. Logo, o cientista de dados jurídico (e o juiz ciborgue) tem um árduo trabalho no exercício da pretensão de correção em relação às próprias técnicas de modelagem de IA judicial a partir dos dados jurídicos.

Para melhor compreensão, veja-se os exemplos gráficos de sub-ajuste e sobreajuste em um modelo de regressão logística:

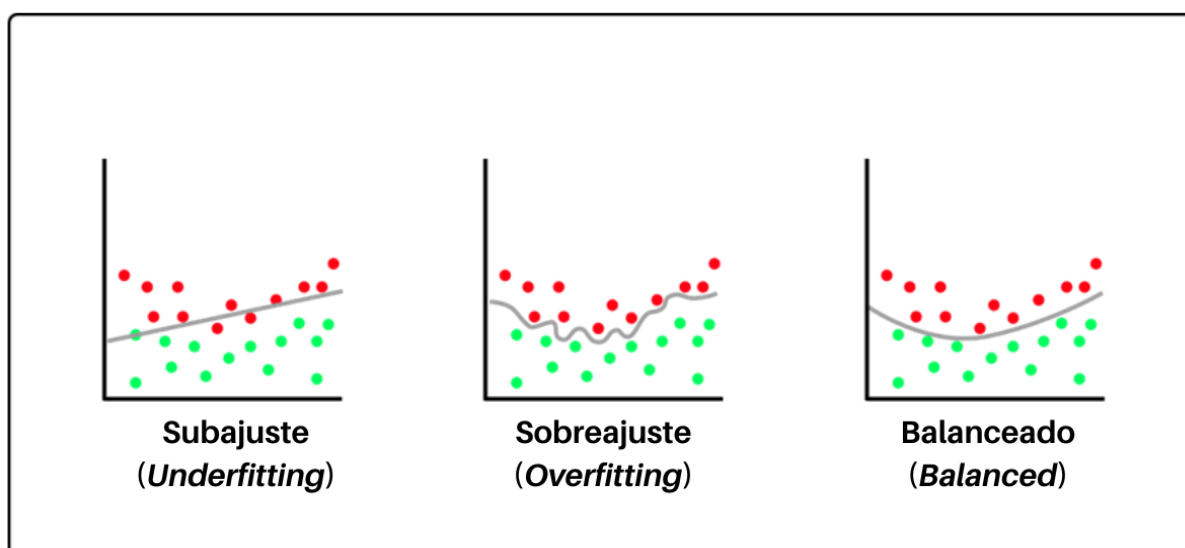


Fig. 10 – Subajuste, sobreajuste e balanceamento em modelo de regressão logística.¹⁶⁵⁷

A próxima fase no ciclo CRISP-DM é a **avaliação**, na qual devem ser estimados os resultados da aplicação do modelo de maneira mais criteriosa, a fim de que se possa obter maior confiabilidade a respeito dos dados e do modelo.¹⁶⁵⁸ Nesse sentido, na fase de avaliação, busca-se analisar se os resultados obtidos a partir do modelo de IA representam regularidades verdadeiras e se tem adequada correlação com os

¹⁶⁵⁷ LIN, David Chuan-En. **8 Simple Techniques to Prevent Overfitting**. [2020]. Disponível em: <https://towardsdatascience.com/8-simple-techniques-to-prevent-overfitting-4d443da2ef7d>. Acesso em: 31 jan. 2023.

¹⁶⁵⁸ PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 31.

dados modelados, para que falsas correlações ou anomalias de amostragem não sejam levadas em consideração no momento da implementação do modelo em um sistema de IA judicial.

Dessa forma, embora seja possível buscar resultados logo após a mineração dos dados, pular a fase de avaliação não é uma boa prática e pode levar às mais diversas consequências danosas. Nessa linha, “geralmente, é muito mais fácil, mais barato, mais rápido e mais seguro testar um modelo primeiro em um ambiente controlado de laboratório”.¹⁶⁵⁹

Nesse sentido, tendo em vista que o objetivo principal da abordagem a partir da Ciência de Dados “é apoiar a tomada de decisão”¹⁶⁶⁰, ou seja, é parte de uma solução maior (consistente outros subsistemas de IA judicial e na decisão final tomada pelo julgador), a fase de avaliação garante que o modelo de aprendizado de máquina satisfaça os objetivos definidos no projeto através da análise de desempenho no mundo real e da probabilidade de existência de erros. Para dar conta desta avaliação qualitativa, o assessor-desenvolvedor (cientista de dados jurídicos, no caso) deve ter como premissa a compreensibilidade do modelo¹⁶⁶¹ (o que é uma faceta do direito fundamental à explicabilidade, nos moldes delineados no subcapítulo 2.6).

Outro desafio é construir um ambiente de produção que reflita a realidade do sistema de IA judicial, a fim de que a avaliação seja a mais abrangente possível¹⁶⁶² antes da implantação do modelo de aprendizado de máquina.

A avaliação de um modelo pode se relacionar a diversos fatores. Contudo, para os limites da presente pesquisa na interação entre Direito e Inteligência Artificial, cabe falar rapidamente dos critérios de **acurácia, sensibilidade, precisão e F1 score**.

A **acurácia** representa o percentual geral de acerto de um modelo de IA. Ou seja, representa a proporção de acertos (verdadeiros positivos e verdadeiros negativos) considerando todas as predições (o que inclui os falsos positivos e os falsos negativos)¹⁶⁶³, conforme a seguinte representação gráfica:

¹⁶⁵⁹ Id.

¹⁶⁶⁰ Id.

¹⁶⁶¹ “E se o modelo em si não é compreensível (por exemplo, talvez o modelo seja uma fórmula matemática muito complexa), como os cientistas de dados podem trabalhar para tornar compreensível o comportamento do modelo?” (Ibid., p. 32).

¹⁶⁶² Id.

¹⁶⁶³ “Accuracy (A): the ratio of correct case predictions over all case predictions. $(TN + TP)/(TN + TP + FN + FP)$ ” (ASHLEY, Kevin D. *Artificial Intelligence and Legal Analytics*, p. 114).

		PREDIÇÃO	
		0	1
OBSERVAÇÃO	0	VN	FP
	1	FN	VP

$$\text{Acurácia} = \frac{VN + VP}{VN + VP + FN + FP}$$

Fig. 11 – Representação da acurácia na matriz da confusão.

Baixos índices de acurácia podem ocorrer pelo fato do conjunto de dados ser desbalanceado, quando existem mais dados de uma determinada classe do que de outra (por exemplo, poucas sentenças de procedência e muitas sentenças de improcedência no *dataset* a ser explorado).

Por outro lado, deve ser sempre lembrado que olhar apenas para a acurácia do modelo pode dar a falsa impressão que a Inteligência Artificial tem um desempenho adequado. Dessa forma, a acurácia não é a única métrica a ser considerada, pois a performance deve ser avaliada no contexto do problema jurídico com um todo.

Já a **sensibilidade** (também conhecida como *recall*¹⁶⁶⁴) representa o percentual de verdadeiros positivos em relação ao total de predições realizadas no conjunto de dados de teste.¹⁶⁶⁵ Veja-se os seguintes gráficos:

		PREDIÇÃO	
		0	1
OBSERVAÇÃO	0	VN	FP
	1	FN	VP

$$\text{Sensibilidade} = \frac{VP}{VP + FN}$$

Fig. 12 – Representação da sensibilidade na matriz da confusão.

¹⁶⁶⁴ “Recall (R): the ratio of positive case predictions that are correct over the number of cases that were positive. (TP)/(TP + FN)” (Id.)

¹⁶⁶⁵ LAGE, Fernanda de Carvalho. Op. cit., p. 308-310.

A fim de ultrapassar a mera trivialidade¹⁶⁶⁶, a sensibilidade deve ser utilizada em conjunto com a acurácia, uma vez que representa a taxa de acerto dos verdadeiros positivos (aquilo que se busca com o modelo). Assim, se errar pode ser catastrófico para a aplicação do modelo de IA no contexto jurídico, a análise da sensibilidade é obrigatória.

Noutro aspecto, a **precisão** (também conhecida como valor preditivo positivo)¹⁶⁶⁷ representa a divisão dos verdadeiros positivos pela soma de todas as predições positivas feitas pelo modelo de IA (verdadeiros positivos e falsos positivos)¹⁶⁶⁸:

		PREDIÇÃO	
		0	1
OBSERVAÇÃO	0	VN	FP
	1	FN	VP

$$\text{Valor Preditivo Positivo} = \frac{VP}{VP + FP}$$

Fig. 13 – Representação da precisão na matriz da confusão.

Portanto, para a aferição da precisão do modelo, são descartadas as predições negativas (verdadeiras e falsas), o que indica o desempenho da IA em casos em que acerta a variável alvo. Em suma, a precisão valoriza os acertos, diminuindo a falsa impressão de segurança que um alto índice de acurácia pode causar.

No mais, o **F1 score**¹⁶⁶⁹ é a média harmônica entre a taxa de precisão e a taxa de sensibilidade¹⁶⁷⁰, o que se apresenta como mais uma forma de se verificar a correção do modelo de IA.

A partir dos índices de acurácia, sensibilidade, precisão e *F1 score*, evidencia-se que o processo de desenvolvimento de modelos de IA baseados em aprendizado

¹⁶⁶⁶ TAULLI, Tom. Op. cit., p. 55.

¹⁶⁶⁷ “Precision (P): the ratio of the number of positive case predictions that are correct over the total number of positive case predictions. $(TP)/(TP + FP)$ ” (ASHLEY, Kevin D. **Artificial Intelligence and Legal Analytics**, p. 114).

¹⁶⁶⁸ LAGE, Fernanda de Carvalho. Op. cit., p. 305-308.

¹⁶⁶⁹ “F1-score or F1-measure: the harmonic mean of precision and recall where both measures are treated as equally important. $2 * (P * R)/(P + R)$ ” (ASHLEY, Kevin D. **Artificial Intelligence and Legal Analytics**, p. 114).

¹⁶⁷⁰ LAGE, Fernanda de Carvalho. Op. cit., p. 310-312.

de máquina conta com a busca pelo acerto, pela melhor resposta a cada caso, de maneira representar **incremento informacional para o apoio à decisão humana**¹⁶⁷¹, o que atua concomitantemente nos campos da correção normativa (a interpretação que mais se adequa ao sistema normativo) e da correção moral (a aplicação do Direito que mais se coaduna com os padrões morais representados nos procedentes ou no sistema normativo). A maior ou menor aderência a uma das espécies de pretensão de correção dependerá do problema jurídico explorado, das premissas utilizadas no processo de desenvolvimento da IA e da própria existência da correção do modelo de aprendizado de máquina. Nesse caminho, ainda que existam percentuais de erro (consideradas acurácia, sensibilidade e precisão), a pretensão de correção é manifestada.

Em resumo, o juiz ciborgue, na perspectiva de um cientista de dados jurídicos que está ciente do duplo canal de correção, deve olhar para o conjunto dos critérios de medição de desempenho do modelo, e não apenas para um dos critérios. O critério mais importante será determinado pelo problema jurídico que se busca solucionar, pois é a partir do problema que se desenvolverá a solução de IA judicial advinda do paradigma *Data Science*.

Noutro aspecto, devem ser utilizadas **técnicas de validação** como *holdout* e validação cruzada K-Fold, por exemplo, para que sejam selecionados os modelos (ou abordagens) com o melhor desempenho dentre as opções disponíveis, bem como para avaliar a performance do modelo escolhido (seu percentual de erro de generalização, conforme os indicadores escolhidos).

A última fase do ciclo CRISP-DM consiste na **implantação**¹⁶⁷², momento em que os subsistemas da IA judicial baseados em aprendizado de máquina serão utilizados em dados do mundo real, ou seja, é quando a IA judicial é colocada em produção.

Neste ponto, um aspecto de destaque a ser deliberado é se deve existir a implantação do modelo de IA judicial em si mesmo ou, num passo menos pretensioso,

¹⁶⁷¹ “Não há dúvidas de que as soluções para essas áreas são relevantes, e a maior sinergia entre IA e Direito, neste contexto, é obtida quando há a diretriz de otimização que a IA pode fazer ao juiz, destacadamente quando se fala em ferramentas flexíveis (pois o julgamento depende desta flexibilidade) de apoio à decisão e de execução de tarefas na elaboração de documentos judiciais. Nesse caminho, a IA pode tornar o trabalho judicial mais confortável, mantendo as exigências de celeridade, qualidade, profundidade e sensibilidade” (HARTMANN PEIXOTO, Fabiano; SILVA, Roberta Zumblick Martins da. Op. cit., p. 122).

¹⁶⁷² PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 32.

se deve ocorrer a implantação de mineração de dados numa perspectiva mais exploratória¹⁶⁷³ do *dataset* (solução típica da jurimetria¹⁶⁷⁴).

A partir dessa visão necessária do ciclo de desenvolvimento de modelos de *machine learning*, fica evidente que não pode existir sistema de IA judicial gestado exclusivamente em gabinete de magistrado ou enclausurado no setor Tecnologia da Informação de qualquer tribunal, pois o isolamento¹⁶⁷⁵ é uma fórmula para o fracasso do projeto de IA judicial. Sistemas de IA judicial devem surgir da conjunção de elementos jurídicos e da tecnologia, motivo pelo qual a pura e simples delegação de estruturação de projetos de IA judicial à equipe de Tecnologia da Informação¹⁶⁷⁶ é uma prática contraproducente que, além de diminuir a legitimidade da futura aplicação do sistema pela redução ou exclusão da participação dos operadores do Direito (magistrados e suas equipes técnicas), revela desperdício de tempo, de recursos financeiros e de trabalho humano.

Destarte, para evitar aquilo que Provost e Fawcett chamam de transferências “em cima do muro” na fase implantação¹⁶⁷⁷, deve existir o acompanhamento perene dos profissionais oriundos do Direito (juízes e assessores-desenvolvedores) em

¹⁶⁷³ “Duas razões principais para a implantação de mineração de dados em si em vez dos modelos produzidos por um sistema de mineração de dados são (i) o mundo pode mudar mais rapidamente do que a equipe de data science pode se adaptar, assim como acontece com detecção de fraude e de invasão e (ii) um negócio tem muitas tarefas de modelagem para que sua equipe de data science depure manualmente cada modelo. Nesses casos, pode ser melhor implantar a fase de mineração de dados na produção. Ao fazê-lo, é fundamental que o processo de instrumentação alerte a equipe de data science sobre qualquer anomalia aparente e forneça uma operação segura” (Ibid., p. 33).

¹⁶⁷⁴ “A jurimetria refuta da visão ‘predeterminista’ do Direito e oferece como alternativa uma visão estocástica. A visão estocástica entende que, para sistemas de maior complexidade, com fatores numerosos e mecanismos de interação complexos, a predeterminação exata dos estados futuros fica inviabilizada” (NUNES, Marcelo Guedes. Op. cit., p. 143).

¹⁶⁷⁵ É necessário que o estranhamento com o diferente, que por vezes surge nas relações ente profissionais do Direito e desenvolvedores, seja transmutado em cooperação ampla e aberta.

¹⁶⁷⁶ “Isso pode ser um erro, porque o processo de mineração de dados é investigativo, aproximando-se mais de pesquisar e desenvolver do que de construir. O ciclo CRISP baseia-se na exploração; ele itera abordagens e estratégias, em vez de projetos de software. Os efeitos são muito menos certos, e os resultados de uma determinada etapa podem alterar a compreensão fundamental do problema. Construir uma solução de mineração de dados, diretamente para implantação, pode ser um compromisso prematuro caro. Em vez disso, projetos de análise devem se preparar para investir em informações que visam reduzir a incerteza de várias maneiras” (PROVOST, Foster; FAWCETT, Tom. Op. cit., p. 34).

¹⁶⁷⁷ “Em termos práticos, há riscos com transferências ‘em cima do muro’ de data science para o desenvolvimento. Pode ser útil lembrar a máxima: ‘Seu modelo não é o que os cientistas de dados projetam, é o que os engenheiros construíram.’ Do ponto de vista de gestão é aconselhável ter membros da equipe de desenvolvimento envolvidos no início do projeto de data science. Eles podem começar como consultores, oferecendo uma visão crítica para a equipe de data science. Esses desenvolvedores assumem gradualmente mais responsabilidade conforme o projeto amadurece. Em algum momento, os desenvolvedores assumirão a liderança e a propriedade do produto. Geralmente os cientistas de dados ainda devem permanecer envolvidos no projeto durante a implantação final, como consultores ou desenvolvedores, dependendo de suas habilidades” (Ibid., p. 33).

relação aos aspectos da implantação dos sistemas de IA judicial, de maneira que o incremento informacional e a legitimidade constitucional inerentes ao duplo canal de correção sejam sempre possibilitados.

No mais, deve ser citado que o ciclo CRISP-DM, devido aos seus aspectos exploratórios, pode ser sempre iniciado novamente, voltando-se às fases anteriores à implantação, para que existam ajustes em relação à definição do problema jurídico ou ainda para a obtenção de novos dados voltados à modelagem do sistema de IA judicial, uma vez que que “um projeto deve ser flexível o suficiente para rever as etapas anteriores com base nas descobertas feitas”.¹⁶⁷⁸

Desse modo, essa foi uma apresentação resumida do ciclo CRISP-DM¹⁶⁷⁹ que levou em conta as peculiaridades de sua aplicação para o desenvolvimento de sistemas de Inteligência Artificial judicial.

5.7 *Deep Learning*: O Aprendizado Profundo e sua Aplicabilidade para o Juiz Ciborgue

As redes neurais de aprendizado profundo merecem um subcapítulo à parte, haja vista que o paradigma *deep learning* possui peculiaridades próprias e apresenta utilidade para o desenvolvimento de sistemas de Inteligência Artificial com aplicabilidade nas rotinas do Poder Judiciário, apesar das dificuldades de sua implantação relativas ao volume e variabilidade dos dados¹⁶⁸⁰, às grandes exigências de desempenho computacional¹⁶⁸¹ e ao uso de bancos de dados relacionais e não relacionais em computação distribuída.¹⁶⁸²

¹⁶⁷⁸ Ibid., p. 34.

¹⁶⁷⁹ Observa-se que questões técnicas mais aprofundadas extrapolariam os limites da presente obra, mas o eventual interesse de aprofundamento pode ser realizado na bibliografia indicada.

¹⁶⁸⁰ LAGE, Fernanda de Carvalho. Op. cit., p. 89-90.

¹⁶⁸¹ Ibid., p. 90.

¹⁶⁸² LOPEZ, Martin Andreoni; SANZ, Igor Jochem; LOBATO, Antonio G. P. Aprendizado de Máquina em Plataformas de Processamento Distribuído de Fluxo: Análise e Detecção de Ameaças em Tempo Real. **Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos (SBRC) - Minicursos**, maio 2018. Disponível em: <http://143.54.25.88/index.php/sbrcminicursos/article/view/1768>. Acesso em 28 nov. 2022.

O uso de algoritmos de redes neurais é tradicionalmente estudado no paradigma *machine learning* (*Perceptron*¹⁶⁸³, *Rede Neural Feed-Forward*¹⁶⁸⁴ e *Rede Neural Backpropagation*¹⁶⁸⁵, por exemplo), o que se inclui no contexto da aprendizagem não supervisionada, tema que, em certa medida, pode ser englobado no âmbito do desenvolvimento estudado no subcapítulo anterior.

Todavia, no paradigma *deep learning* há o aumento da interconexão na rede neural, a fim de enfrentar desafios computacionais mais complexos, assim como os “neurônios reais, neurônios artificiais começam a ficar mais interessantes quando você começa a conectá-los”.¹⁶⁸⁶ Para facilitar a compreensão, segue exemplo de representação gráfica de uma rede neural artificial:

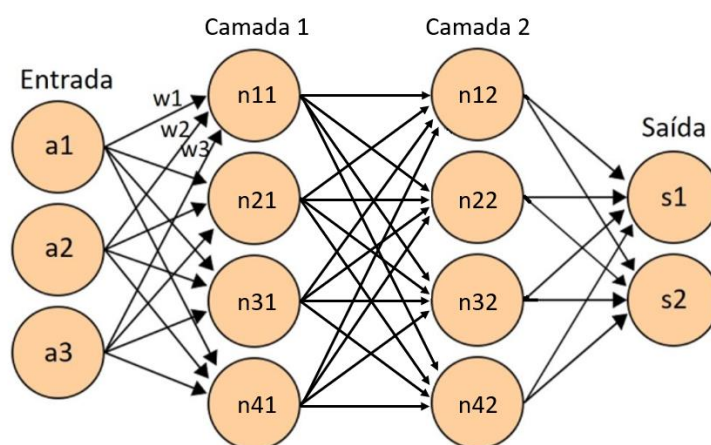


Fig. 14 – Representação simples de uma rede neural, seus neurônios artificiais e suas camadas.¹⁶⁸⁷

¹⁶⁸³ “A rede neural mais simples é a *perceptron*, que aproxima um único neurônio com n entradas binárias. Ela computa a soma ponderada de entradas e ‘dispara’ se essa soma for maior que zero” (GRUS, Joel. Op. cit., p. 213).

¹⁶⁸⁴ “A topologia do cérebro é demasiadamente complicada, então é normal aproximá-la com uma rede neural *feed-forward* idealizada que consiste de camadas discretas de neurônios, cada uma conectada à seguinte. Isso tipicamente envolve uma camada de entrada (que recebe entradas e as transmite sem modificações), uma ou mais ‘camadas ocultas’ (em que cada uma consiste de neurônios que pegam saídas da camada anterior, fazem um cálculo e passam o resultado para a próxima camada), e uma camada de saída (que produz as saídas finais).

Assim como o perceptron, cada neurônio (não de entrada) possui o peso correspondente a cada uma de suas entradas e uma polarização (tendência). Para simplificar nossa representação, adicionaremos a polarização (*bias*) no final do vetor de pesos e daremos a cada neurônio uma *entrada polarizada* que é sempre igual a 1” (Ibid., p. 215).

¹⁶⁸⁵ “Geralmente nós não construímos redes neurais manualmente. Isso se dá, em parte, porque as usamos para resolver problemas muito maiores – um problema de reconhecimento da imagem pode envolver dezenas ou milhares de neurônios. E em parte porque nós geralmente não conseguimos ‘raciocinar’ sobre o que neurônios deveriam ser.

Em vez disso, nós usamos dados para treinar redes neurais. Uma abordagem popular é um algoritmo chamado *backpropagation* que possui semelhanças com o algoritmo de gradiente descendente que vimos anteriormente” (Ibid., p. 218).

¹⁶⁸⁶ Ibid., p. 215.

¹⁶⁸⁷ DIDÁTICA Tech. **O que são Redes Neurais e Deep Learning**. [2022]. Disponível em: <https://didatica.tech/introducao-a-redes-neurais-e-deep-learning>. Acesso em: 31 jan. 2023.

Sem a necessidade de aprofundamento nos aspectos técnicos de programação, matemática e estatística envolvidos, tendo por base a abordagem jurídica desta obra, cabe referir que alguns dos principais exemplos de algoritmos de aprendizagem de máquina profunda são: a) Rede Neural Recorrente – *RNN*¹⁶⁸⁸; b) Rede Neural Convolucional – *CNN*¹⁶⁸⁹; c) Redes Adversárias Generativas – *GAN*¹⁶⁹⁰; e d) Rede Neural *Transformer*.¹⁶⁹¹

Os modelos de Inteligência Artificial baseados em *deep learning* são desenvolvidos através do envio de sinais por meio de neurônios artificiais (algoritmos que funcionam como portas lógicas que imitam o funcionamento dos neurônios biológicos), os sinais são transmitidos de um neurônio artificial ao outro (mimetizando as sinapses biológicas), entre cada uma das camadas (*layers*), o que leva à aprendizagem mediante o ajuste dos pesos em cada neurônio por onde passa (que aumenta ou reduz a intensidade do sinal e calibra a importância das conexões entre

¹⁶⁸⁸ “Com uma rede neural recorrente (*Recurrent Neural Network* – *RNN*), a função não só processa a entrada atual, mas também as entradas anteriores ao longo do tempo. Um exemplo disso é o que acontece quando você digita caracteres em um aplicativo de mensagens. À medida que ocorre a digitação, o sistema tenta prever palavras. Então, se você digitar ‘Ele’, o computador vai sugerir ‘Ele’, ‘Elefante’ e ‘Elétrico’, por exemplo. A *RNN* é essencialmente uma sequência de redes neurais que se alimentam umas das outras com base em algoritmos complexos” (TAULLI, Tom. Op. cit., p. 103).

¹⁶⁸⁹ “Intuitivamente, faz sentido ter todas as unidades de uma rede neural conectadas. Isso funciona bem com muitas aplicações.

No entanto, há cenários muito distantes do ideal, como é o caso do reconhecimento de imagens. Basta imaginar como seria complexo um modelo no qual cada pixel é uma unidade! Ele rapidamente se tornaria incontornável. Haveria também outras complicações como a superadaptação (*overfitting*). Isso ocorre quando os dados não refletem o que está sendo testado ou o foco está nas características erradas.

Para lidar com tudo isso, é possível usar uma rede neural convolucional (*Convolutional Neural Network*).” (...)

“Cada um dos estágios é chamado de convolução. Por exemplo, o primeiro nível seria identificar linhas e ângulos; em seguida, o córtex visual encontraria as formas e, então, detectaria os objetos.

Isso é análogo a como uma *CNN* baseada em computador trabalha. Considere o seguinte exemplo: suponha que você queira construir um modelo capaz de identificar uma carta. A entrada da *CNN* será uma imagem com 3.072 pixels. Cada um dos pixels será um valor que vai de 0 a 255, o que indica a intensidade geral. Usando uma *CNN*, o computador passará por diversas variações para identificar os recursos” (Ibid., p. 105).

¹⁶⁹⁰ “Observe o fluxo de trabalho básico:

- Gerador: essa rede neural produz uma infinidade de novas criações, como fotos ou frases.
- Discriminatório: essa rede neural analisa as criações para ver quais são reais.
- Ajustes: com os dois resultados, um novo modelo altera as criações para torná-las o mais realista possível. Por meio de muitas iterações, o discriminatório não precisa mais ser usado.” (...)

“Infelizmente, quando as *GAN*’s são o tema, alguns de seus usos têm sido menos admiráveis. Um exemplo é usá-las para *deepfakes*, que envolvem o uso de redes neurais para criar imagens ou vídeos enganosos” (Ibid., p. 106-107).

¹⁶⁹¹ LI, N.; LIU, S.; LIU, Y.; ZHAO, S.; LIU, M. Neural Speech Synthesis with Transformer Network. **Proceedings of the AAAI Conference on Artificial Intelligence**, v. 33, n 1, p. 6706-6713, 2019. DOI: 10.1609/aaai.v33i01.33016706.

os neurônios), chegando a uma saída (*output*) ao final do processo.¹⁶⁹² Ademais, uma rede neural de aprendizado profundo é dividida em diversas camadas de neurônios que ficam organizadas em camada de entrada (*input layer*), camadas ocultas (*hidden layers*) e a camada de saída (*output layer*).¹⁶⁹³

Assim, um algoritmo de *deep learning* analisa todos dados que tem acesso (*input*) e, “em seguida, encontra as relações usando uma rede neural que imita o cérebro humano”¹⁶⁹⁴, o que pode fazê-lo com pouco ou nenhum pré-processamento de dados. Isso ocorre porque os algoritmos de *deep learning* são mais apropriados para grandes volumes de dados não estruturados¹⁶⁹⁵, sendo que sua atuação não exige rotulagem de dados ou extração de recursos¹⁶⁹⁶, o que representa uma de suas principais vantagens em relação a outras abordagens de IA.

Contudo, ainda que tenha vantagens em determinados cenários, o uso de *deep learning* no desenvolvimento Inteligência Artificial judicial é uma opção a ser avaliada com cautela em cada caso, pois existem desvantagens que, resumidamente, são: a) a maior opacidade¹⁶⁹⁷ do processo de aprendizado da IA (tema já relatado no subcapítulo 2.6); b) a eficiência do treinamento depende de bases de dados muito grandes (o que nem sempre está disponível); e c) a necessidade de grande poder computacional para o processamento (a depender do escopo do projeto de IA).¹⁶⁹⁸ Além disso, a grande maioria dos *datasets* existentes nos órgãos do Poder Judiciário possui algum nível de organização, podendo ser classificados em dados estruturados

¹⁶⁹² “Os modelos de redes neurais de *deep learning* processam sinais enviando-os através de uma rede de nós análogos aos neurônios. Os sinais passam de nó em nó ao longo de links, análogos das junções sinápticas entre neurônios. A aprendizagem melhora o resultado (*output*) por meio do ajuste dos pesos que amplificam ou amortecem os sinais que cada ligação transmite. Os nós são tipicamente organizados em uma série de camadas que são aproximadamente análogas a diferentes centros de processamento no córtex” (LAGE, Fernanda de Carvalho. Op. cit., p. 88).

¹⁶⁹³ “O aprendizado profundo usa uma rede neural para reproduzir a inteligência humana, e existem três tipos de camadas de neurônios em uma rede neural: a camada de entrada (*input layer*), a(s) camada(s) oculta(s) (*hidden layers*) e a camada de saída (*output layer*). As conexões entre neurônios têm um peso determinado pela importância do elemento de entrada. Vale destacar que para treinar uma rede neural profunda, é necessário um conjunto de dados muito grande. Por meio da comparação dos resultados de saída de forma iterativa com os dados incluídos no conjunto será possível calcular uma função de perda indicando quanto o algoritmo está errado. E, após cada iteração, os pesos entre os neurônios são redistribuídos usando o método de gradiente descendente para minimizar a função de perda” (Ibid., p. 89).

¹⁶⁹⁴ TAULLI, Tom. Op. cit., p. 55.

¹⁶⁹⁵ Embora também tenham aplicabilidade em dados rotulados (HARTMANN PEIXOTO, Fabiano; SILVA, Roberta Zumblick Martins da. Op. cit., p. 100).

¹⁶⁹⁶ TAULLI, Tom. Op. cit., p. 55.

¹⁶⁹⁷ GUIDOTTI, Riccardo; MONREALE, Anna; RUGGIERI, Salvatore; TURINI, Franco; GIANNOTTI, Fosca; PEDRESCHI, Dino. Op. cit.

¹⁶⁹⁸ LAGE, Fernanda de Carvalho. Op. cit., p. 90.

ou dados semiestruturados, espécies de base de dados que podem ser tratadas com outras abordagens de IA e Ciência de Dados, a fim de que levem a modelos de *machine learning* com desempenho adequado.

Vale destacar que, embora o paradigma *deep learning* ainda esteja em um estágio inicial de desenvolvimento científico¹⁶⁹⁹ e tenha muito a evoluir na Ciência da Computação¹⁷⁰⁰, a abordagem de aprendizado profundo já tem se mostrado útil no Direito, como, por exemplo, para o reconhecimento de entidades nomeadas (*Named Entity Recognition – NER*) em documentos jurídicos, como pessoas, jurisprudência, legislação, tempo, localização e organização¹⁷⁰¹, o que, por sua vez, apresenta-se como uma tecnologia necessária para o Processamento de Linguagem Natural (PLN) voltado aos textos jurídicos. Além disso, atualmente, tem-se notícia de iniciativa do Tribunal de Justiça do Estado de São Paulo (TJSP) para o desenvolvimento de sistema de IA judicial voltado à extração de informações de guias de recolhimento de custas, projeto que utiliza o paradigma *deep learning* através de Redes Neurais Convolucionais baseadas nas bibliotecas *Keras* e *Tensorflow*.¹⁷⁰²

Ainda que a abordagem *deep learning* possa se mostrar útil para as hipóteses de dados não estruturados e semiestruturados, como no caso de dados de processos judiciais que originalmente tramitavam em forma física (processos impressos em papel) e que foram digitalizados com baixa qualidade de imagem ou sem padronização adequada, há relatos de diversas dificuldades na utilização da abordagem *machine learning* em tais espécies de *datasets*.¹⁷⁰³

Por outro lado, o caráter incipiente do desenvolvimento de sistemas de IA baseados em *deep learning* não afasta, ao menos inicialmente, a hipótese do Judiciário se tornar cliente de sistemas de IA desenvolvidos com o aprendizado de máquina profundo. Por exemplo, há potencial de uso do *GPT-4*¹⁷⁰⁴ da *OpenAI* (ou de

¹⁶⁹⁹ TAULLI, Tom. Op. cit., p. 115.

¹⁷⁰⁰ O que depende também da evolução do poder de processamento dos computadores, remetendo aspectos do problema para a Engenharia da Computação voltada à pesquisa de *chips* de processamento e de novas tecnologias de *hardware*.

¹⁷⁰¹ MOTA, Caio et al. Named Entity Recognition in Portuguese Legal Documents Using Deep Learning. In: **Anais do XVIII Encontro Nacional de Inteligência Artificial e Computacional**, Porto Alegre, p. 130-140, 2021. DOI: 10.5753/eniac.2021.18247.

¹⁷⁰² SALOMÃO, Luis Felipe (coord.) et al. Op. cit., p. 182-184.

¹⁷⁰³ Por exemplo, a equipe de desenvolvimento do Tribunal de Justiça do Estado do Paraná (TJPR) relatou as dificuldades de se trabalhar com dados e extrair informações de processos físicos digitalizados relacionadas às execuções fiscais (PARANÁ. Tribunal de Justiça do Estado do Paraná. **Pedido de Providências 0103379-12.2022.8.16.6000**. Reunião TJPR-UnB realizada em 18/08/2022 (Vídeo nº 8063985 e Ata nº 8067385). Curitiba, PR: Tribunal de Justiça do Estado do Paraná).

¹⁷⁰⁴ OPENAI. **OpenAI API**.

novas versões subsequentes) para se aproveitar de sua relativa eficiência para geração de textos em linguagem natural¹⁷⁰⁵ ou da Inteligência Artificial da Gato da *DeepMind*¹⁷⁰⁶ que tem capacidade multitarefas. Em relação à cautela necessária ao se adotar esta opção (em parte já referida nesta obra), destaca-se que nenhum sistema, próprio do Judiciário ou fornecido por empresa terceira, deve ser colocado em produção sem os devidos cuidados e testes (e aqui remete-se naturalmente ao subcapítulo anterior), pois, por exemplo, o sistema *Galactica* da empresa *Meta*¹⁷⁰⁷ inicialmente demonstrou grande capacidade para resumir artigos científicos (sobre Matemática e Ciências) e para responder perguntas sobre os temas, mas teve de ser interrompido ao se descobrir que estava gerando grande quantidade de informações falsas e enganosas (incluindo citações de trabalhos de pesquisa que nunca foram escritos pelos autores citados).¹⁷⁰⁸ É evidente que, na aplicação do Direito, falhas desta magnitude poderiam ser catastróficas.

Para além da necessária cautela no desenvolvimento da IA, o uso de sistemas terceiros desenvolvidos sob o paradigma *deep learning* pode ser demandado pelos órgãos do Poder Judiciário, desde que haja o devido cuidado no tratamento da segurança da informação, especialmente de dados pessoais.

Nesse caminho, um sistema de Inteligência Artificial judicial pode, quando considerados todos os subsistemas que o integram, ter a ele conectadas (via *API*) soluções de IA através de sistemas de empresas terceiras que queiram prestar serviços ao Poder Judiciário. Destaca-se, uma vez mais, que todo cuidado é necessário e que o tema merece bastante estudo científico e grande debate democrático sobre os riscos e potencialidades do uso de *deep learning* no Judiciário brasileiro, o que, noutro foco, não pode ser um entrave total.

Logo, desde que ponderados os riscos (respeito ao direito fundamental à explicabilidade, por exemplo – vide o subcapítulo 2.5), o potencial de utilização das tecnologias *deep learning* no Direito é bastante evidente, o que não pode ser negligenciado pela doutrina que queira tratar a questão com seriedade.

¹⁷⁰⁵ BROWN, Tom B. et al. Op. cit.

¹⁷⁰⁶ REED, Scott et al. A Generalist Agent. **Transactions on Machine Learning Research**. [2022]. Disponível em: <https://openreview.net/pdf?id=1ikK0kHvj>. Acesso em: 16 nov. 2022.

¹⁷⁰⁷ TAYLOR, Ross et al. **Galactica**: A Large Language Model for Science. [2022]. Disponível em: <https://galactica.org/static/paper.pdf>. Acesso em: 16 nov. 2022.

¹⁷⁰⁸ ROSE, Janus. **Facebook Pulls Its New ‘AI For Science’ Because It’s Broken and Terrible**. [2022]. Disponível em: <https://www.vice.com/en/article/3adyw9/facebook-pulls-its-new-ai-for-science-because-its-broken-and-terrible>. Acesso em 29 nov. 2022.

5.8 Processamento de Linguagem Natural: Uma Conversa entre a IA e o Juiz Ciborgue

O processamento de linguagem natural (PLN), ou *Natural Language Processing (NLP)*, consiste no tratamento da linguagem humana através do processamento computacional, de maneira que o sistema carregue os dados linguísticos (por escrito ou por voz) através do reconhecimento de entidades¹⁷⁰⁹ (nomes de pessoas, lugares, etc.), compreenda (ainda que parcialmente) o que foi dito e, a depender de seu objetivo, estabeleça comunicação com o ser humano.¹⁷¹⁰

Para a compreensão da complexidade envolvida no Processamento de Linguagem Natural (PLN), a linguagem de máquina é baseada em lógica e matemática¹⁷¹¹, sendo que a linguagem humana, ainda que também tenha sua lógica, possui contexto e é baseada em abstrações que são compreensíveis apenas aos falantes que participam da conversação.¹⁷¹² A partir disso, verifica-se que existem desafios¹⁷¹³ para o desenvolvimento de sistemas de IA judicial baseados em processamento de linguagem natural, sendo os principais: a) a ambiguidade e a vagueza das palavras; b) a linguagem se modifica conforme muda o mundo que ela representa; c) a fala coloquial do ser humano contém erros gramaticais; d) existe a variabilidade decorrente de sotaques e dialetos; e) o significado das palavras pode ser influenciado pelo humor do falante e pelo contexto; f) muitas palavras têm o mesmo significado (porém com nuances); g) muitas palavras podem ter diversos significados (polissemia); h) as conversas humanas podem conter interrupções (ser não lineares); i) linguagem implícita; j) linguagem criada para diversão; e k) mudanças decorrentes do contexto e da pessoa que fala.¹⁷¹⁴

Nesse sentido, a respeito do Processamento de Linguagem Natural, Stuart Russell e Peter Norvig destacam que um “agente que deseja **adquirir conhecimento** precisa entender (pelo menos parcialmente) a ambígua e confusa linguagem que os

¹⁷⁰⁹ FERNANDES, Ivo André Domingues. **A Deep Learning Approach to Named Entity Recognition in Portuguese Texts**. 2010. Dissertação (Mestrado Integrado em Engenharia Informática e Computação). Faculdade de Engenharia da Universidade do Porto (FEUP), Porto, 2018.

¹⁷¹⁰ RUSSELL, Stuart; NORVIG, Peter. Op. cit., p. 751.

¹⁷¹¹ GRUS, Joel. Op. cit., 213-224.

¹⁷¹² HABERMAS, Jürgen. **Teoria do Agir Comunicativo**, I, p. 40.

¹⁷¹³ TAULLI, Tom. Op. cit., p. 136-137.

¹⁷¹⁴ Por exemplo, a frase “está chovendo” pode representar uma descrição (o evento chuva) ou pode ser uma ordem de uma mãe para que seu filho pegue o casaco ou o guarda-chuva.

seus humanos usam”.¹⁷¹⁵ O caminho para o referido objetivo passa pelas fases de classificação de texto, recuperação de informação e extração de informação¹⁷¹⁶, o que, na presente obra, foi tratado no subcapítulo 5.6, sendo desnecessária a repetição do tema.

Conforme dados do ano de 2022, levantados pela Fundação Getúlio Vargas, no Judiciário brasileiro os principais projetos de IA que usam técnicas de Processamento de Linguagem Natural (PLN) são: a) *Identificação de Fundamentos de Inadmissão do REsp*¹⁷¹⁷, do Superior Tribunal de Justiça; b) *Análise Legal Inteligente – ALEI*¹⁷¹⁸, do Tribunal Regional Federal da 1ª Região; c) *Agrupamento de Apelações por Similaridade de Sentença*¹⁷¹⁹, do Tribunal Regional Federal da 4ª Região; d) *I.A. de Classificação de Petições Intermediárias*, do Tribunal de Justiça do Estado do Acre¹⁷²⁰; e) *Robô Hércules*¹⁷²¹, do Tribunal de Justiça do Estado de Alagoas; f) sistema o *TIA*¹⁷²², do Tribunal de Justiça do Estado de Amapá; g) *Programa Cientista Chefe*¹⁷²³, do Tribunal de Justiça do Estado de Ceará; h) *BERNA – Busca Eletrônica em Registros usando Linguagem Natural*¹⁷²⁴, do Tribunal de Justiça do Estado de Goiás; i) *Jurimetria com a IA*¹⁷²⁵, do Tribunal de Justiça do Mato Grosso do Sul; j) *MIDAS*¹⁷²⁶, do Tribunal de Justiça do Estado da Paraíba; k) *Peticionamento Inteligente*¹⁷²⁷, do Tribunal de Justiça do Estado de Rondônia; l) *IA Execução Fiscal*¹⁷²⁸, do Tribunal de Justiça do Estado do Rio Grande do Sul; m) *Chatbot DIGEP*¹⁷²⁹, também do TJRS; n) *Classificador de petições em Execuções Fiscais*¹⁷³⁰, do Tribunal de Justiça do Estado de Santa Catarina; o) *JUDI (Chatbot)*¹⁷³¹, do Tribunal de Justiça do Estado de São Paulo; p) projeto *Inteligência Artificial como ferramenta*

¹⁷¹⁵ RUSSELL, Stuart; NORVIG, Peter. Op. cit., p. 751.

¹⁷¹⁶ LAGE, Fernanda de Carvalho. Op. cit., p. 273-282.

¹⁷¹⁷ SALOMÃO, Luis Felipe (coord.) et al. Op. cit., p. 68-70.

¹⁷¹⁸ Ibid., p. 75-77.

¹⁷¹⁹ Ibid., p. 81-83.

¹⁷²⁰ Ibid., p. 87-89.

¹⁷²¹ Ibid., p. 90-92.

¹⁷²² Ibid., p. 93-95.

¹⁷²³ Ibid., p. 102-104.

¹⁷²⁴ Ibid., p. 123-125.

¹⁷²⁵ Ibid., p. 132-134.

¹⁷²⁶ Ibid., p. 142-144.

¹⁷²⁷ Ibid., p. 161-163.

¹⁷²⁸ Ibid., p. 167-169.

¹⁷²⁹ Ibid., p. 170-172.

¹⁷³⁰ Ibid., p. 179-181.

¹⁷³¹ Ibid., p. 185-187.

de auxílio jurisdicional¹⁷³², do Tribunal de Justiça do Estado de Sergipe; q) *Projeto Gemini*¹⁷³³, dos Tribunais Regionais do Trabalho da 5ª, da 15ª e da 17ª Regiões; r) sistema, do Tribunal de Justiça do Estado do Pará¹⁷³⁴, voltado à migração de dados de processos físicos para a base de dados de eletrônica; e, s) também no TJPA, um sistema para auxílio na movimentação processual.¹⁷³⁵ Logo, observa-se que as técnicas de Processamento de Linguagem Natural (PLN) têm sido bastante utilizadas para o desenvolvimento de sistemas de IA judicial no Brasil.

Percebe-se que o Processamento de Linguagem Natural (PLN) pode se utilizar de técnicas mais clássicas de desenvolvimento de IA (como lógica de bayesiana e cadeias de Markov)¹⁷³⁶, até os paradigmas *machine learning*¹⁷³⁷ e *deep learning*.¹⁷³⁸

A depender do modelo de IA a ser adotado no Processamento de Linguagem Natural, pode existir maior ou menor trabalho humano na fase de preparação dos dados, sendo que para o PLN são necessárias técnicas de: **a) tokenização** (já referida no subcapítulo 5.6), que ocorre por meio da eliminação da pontuação e de caracteres especiais (de maneira que cada palavra corresponda um único *token*), da junção de palavras (n-gramas), para possam ser percebidas como uma única palavra, e da eliminação de *stopwords* (palavras com alta taxa de reiteração no texto, mas que têm pouco conteúdo informativo¹⁷³⁹); **b) estemização**¹⁷⁴⁰ mediante eliminação de afixos e sufixos para a redução das palavras às suas raízes¹⁷⁴¹; e **c) lemanização**¹⁷⁴² através da busca de palavras com significados semelhantes ou muito próximos¹⁷⁴³, procedimento que possibilita que o sistema baseado em PLN possa “entender os significados das palavras e do contexto”¹⁷⁴⁴ (o que exige bastante poder computacional).

¹⁷³² Ibid., p. 188-190.

¹⁷³³ Ibid., p. 215-217.

¹⁷³⁴ Ibid., p. 135-137.

¹⁷³⁵ Ibid., p. 138-140.

¹⁷³⁶ RUSSELL, Stuart; NORVIG, Peter. Op. cit., p. 753.

¹⁷³⁷ GRUS, Joel. Op. cit., p. 241-253.

¹⁷³⁸ TAULLI, Tom. Op. cit., p. 137.

¹⁷³⁹ Por exemplo, os artigos “a”, “o”, “as” e “os”.

¹⁷⁴⁰ TAULLI, Tom. Op. cit., p. 140-141.

¹⁷⁴¹ Por exemplo, as palavras “consultor”, “consultoria” e “consultando” têm como raiz a palavra “consulta”.

¹⁷⁴² TAULLI, Tom. Op. cit., p. 141-142.

¹⁷⁴³ Por exemplo, as palavras “equivalente”, “análogo”, “correspondente” e “similar” possuem significado bastante próximo.

¹⁷⁴⁴ TAULLI, Tom. Op. cit., p. 142.

Em relação aos algoritmos necessários para o Processamento de Linguagem Natural (PLN), esses podem ser variados nos espectros de aprendizagem supervisionada e aprendizagem não supervisionada, motivo pelo qual, em respeito ao escopo precipuamente jurídico desta obra, não serão tratados de algoritmos específicos.

No entanto, deve ser mencionado que os grandes avanços em Processamento de Linguagem Natural ocorreram a partir a adoção do paradigma *deep learning* (vide subcapítulo 5.7), o que merece um destaque maior no presente estudo.

Por exemplo¹⁷⁴⁵, o *ChatGPT* (versão da Inteligência Artificial *GPT-3.5*¹⁷⁴⁶ mais intuitiva e com maior acessibilidade para conversas com humanos) assim respondeu ao lhe ser solicitado para explicar a Teoria da Justiça em poucas palavras:

“A teoria da justiça é um campo da filosofia política que se ocupa de analisar e estabelecer os critérios pelos quais se pode distinguir o que é justo ou injusto em uma sociedade. Há muitas teorias diferentes sobre o que constitui a justiça, cada uma delas com suas próprias premissas e conclusões. Algumas teorias enfatizam a igualdade entre os indivíduos, enquanto outras defendem a liberdade individual como o principal critério de justiça. Algumas teorias também podem incluir outros elementos, como a distribuição de recursos ou a proteção dos direitos humanos. A teoria da justiça é um tópico importante para os filósofos políticos e juristas, pois ajuda a estabelecer os padrões éticos e normativos que devem reger as relações sociais e políticas em uma sociedade”.¹⁷⁴⁷

Em língua inglesa a resposta do *ChatGPT* para a mesma pergunta foi um pouco diferente, mas manteve o contexto.¹⁷⁴⁸ Todavia, adverte-se que a variabilidade

¹⁷⁴⁵ Destaque-se que **não é objeto da presente pesquisa fazer uma análise exaustiva da inteligência artificial *GPT-3* da *OpenAI*** (o que, provavelmente, demandaria uma tese própria em Ciência da Computação ou em Linguística), sendo que a referência ao referido sistema computacional em certas passagens desta obra (e em especial neste subcapítulo) se dá em um cunho exploratório e pedagógico, a fim de que uma das maiores inteligências artificiais já criadas pela humanidade possa ser exemplificada em seus limites e possibilidades de aplicabilidade ao Direito, o que facilita a compreensão de conceitos como Processamento de Linguagem Natural (PLN) e *deep learning*, dentre outros.

¹⁷⁴⁶ A inteligência artificial *GPT-4* já foi relatada ao longo da presente obra. Para maiores informações vide os subcapítulos 2.1 e 5.7. Os detalhes técnicos do desenvolvimento da versão anterior (*GPT-3*) estão descritos em: BROWN, Tom B. et al. Op. cit.

¹⁷⁴⁷ CHATGPT. **Teoria da Justiça em poucas palavras**. Disponível em: <https://chat.openai.com/chat>. Acesso em 20 dez. 2022.

¹⁷⁴⁸ “*The theory of justice is a concept in political philosophy that refers to the fair and just distribution of goods and burdens in a society. The theory of justice is typically concerned with questions of equality and fairness, and it often involves the notion that all individuals should be treated equally under the law.*”

inerente ao caráter estocástico dos textos produzidos pelo *GPT-3.5* e pelo *GPT-4*, e o aprendizado da IA, pode levar a respostas um pouco diferentes em cada caso. Na sequência, foi feita a seguinte pergunta ao *ChatGPT*: “Você conhece Direito Constitucional do Brasil?” A Inteligência Artificial assim respondeu:

“Sim, eu posso ajudar com questões sobre Direito Constitucional do Brasil. O Direito Constitucional é o ramo do direito que trata da Constituição de um país, que é o documento mais importante de um Estado e estabelece os direitos e deveres dos cidadãos, os poderes do governo e a organização política e jurídica do país. A Constituição Brasileira é a lei fundamental do Brasil e foi promulgada em 1988. Ela estabelece os direitos e garantias fundamentais dos cidadãos, os poderes e deveres do Estado, a organização dos poderes Executivo, Legislativo e Judiciário, entre outras coisas. Se você tiver alguma pergunta específica sobre o Direito Constitucional do Brasil, fique à vontade para me perguntar”.¹⁷⁴⁹

Diante da abertura para novas perguntas, foi feito o seguinte questionamento ao *ChatGPT*: “Ok, pode me explicar a teoria da simetria constitucional?” E a Inteligência Artificial assim respondeu:

“A teoria da simetria constitucional é uma teoria jurídica que defende a existência de uma simetria entre os poderes do Estado e as garantias constitucionais dos cidadãos. Segundo essa teoria, os poderes do Estado, como o Executivo, o Legislativo e o Judiciário, devem ser equilibrados e exercidos de forma simétrica em relação aos direitos e garantias fundamentais dos cidadãos.

Essa teoria surgiu como uma forma de proteger os direitos fundamentais dos cidadãos e garantir o devido processo legal em situações em que o Estado exerce seus poderes. Ela também é vista como uma forma de evitar o abuso

One of the most influential theories of justice is known as distributive justice, which proposes that the distribution of goods and burdens in a society should be based on principles of fairness and equality. This could involve, for example, ensuring that everyone has access to basic necessities such as food, shelter, and healthcare. Other theories of justice may focus on other aspects of fairness and equality, such as procedural justice, which emphasizes the importance of fair and unbiased processes in decision-making, or restorative justice, which focuses on repairing the harm caused by criminal behavior” (CHATGPT. **Theory of Justice in simple terms.** Disponível em: <https://chat.openai.com/chat>. Acesso em 20 dez. 2022).

¹⁷⁴⁹ CHATGPT. **Resposta à pergunta: Você conhece Direito Constitucional do Brasil?** Disponível em: <https://chat.openai.com/chat>. Acesso em 20 dez. 2022.

de poder por parte do Estado e assegurar que os direitos dos cidadãos sejam respeitados e protegidos.

A teoria da simetria constitucional é uma das teorias que fundamentam o sistema de freios e contrapesos do Estado de direito, que busca garantir que nenhum dos poderes do Estado tenha um poder absoluto sobre os outros e evitar o excesso de poder por parte de qualquer um deles”.¹⁷⁵⁰

Como pode ser visto, um erro conceitual apareceu, uma vez que a simetria constitucional (art. 25, *caput*, da Constituição Federal), ainda que tenha sua extensão disputada na doutrina e na jurisprudência do Supremo Tribunal Federal, consistente na reprodução obrigatória pelos Estados-membros dos princípios de organização dos Poderes delineados para a União, não se confunde com a separação harmônica entre os Poderes (art. 2º, *caput*, da Constituição Federal).¹⁷⁵¹ Quanto mais este autor perguntou, mais o *ChatGPT* tentou sutilmente sustentar que a sua resposta era a correta. Para tentar superar o entrave, foi fornecida uma síntese do conceito correto (oportunidade em que a IA se desculpou pelo equívoco). Contudo, após bastante processamento, o sistema apresentou erro¹⁷⁵² quando perguntado se responderia corretamente ou insistiria na resposta equivocada caso outra pessoa fizesse a mesma pergunta.

Em relação aos riscos do *ChatGPT* (nas versões *GPT-3.5* e *GPT-4*), a própria *OpenAI* deixa claro que embora “tenha salvaguardas em vigor, o sistema pode ocasionalmente gerar informações incorretas ou enganosas e produzir conteúdo ofensivo ou tendencioso”¹⁷⁵³, motivo pelo qual o mesmo não se destina a dar conselhos.¹⁷⁵⁴

Além disso, mesmo em língua inglesa, o *GPT-3.5* e o *GPT-4* (versões puras da IA e sem os controles inseridos no *ChatGPT*) podem ser psicologicamente

¹⁷⁵⁰ CHATGPT. **Resposta à pergunta: Você pode explicar a teoria da simetria constitucional?** Disponível em: <https://chat.openai.com/chat>. Acesso em 20 dez. 2022.

¹⁷⁵¹ O próprio conceito de separação dos Poderes que permeou a resposta foi bastante tímido, o que também revela viés no treinamento da IA (que é conhecido e relatado nos textos técnicos elaborados pela equipe da *OpenAI*, haja vista que o escopo do projeto não é voltado para o Direito).

¹⁷⁵² “*That model is currently overloaded with other requests. You can retry your request, or contact us through our help center at help.openai.com if the error persists. (Please include the request ID 14bd5a3d9a2978ff471125fbb6f5779b in your message.)*” (OPENAI. **ChatGPT**. [2022]. Disponível em: <https://chat.openai.com/chat>. Acesso em 20 dez. 2022).

¹⁷⁵³ Tradução livre de: “*While we have safeguards in place, the system may occasionally generate incorrect or misleading information and produce offensive or biased content. It is not intended to give advice*” (Id.).

¹⁷⁵⁴ Id.

envolventes para quem não tenha capacidades de os desafiar, já em português mostraram menor criatividade (ainda que o *GPT-4* tenha se saído melhor), conforme testes realizados por este autor na busca de encontrar características para além do descrito em textos técnicos de Ciência da Computação.

Destaca-se que, desde de maio de 2021, este autor tem reiteradamente testado a Inteligência Artificial *GPT-3* da *OpenAI*¹⁷⁵⁵, tendo posteriormente continuado os testes no *GPT-3.5* e no *GPT-4*. O comportamento apresentado pelo *ChatGPT* em março de 2023 representa a ausência de correção em Direito que foi verificada em todo o período de testes que este autor empreendeu.¹⁷⁵⁶ Ainda que se trate de uma IA “*task agnostic*”, naturalmente não se espera que o *GPT-3.5* e o *GPT-4* tenham eficiência em Direito, uma vez que não foram adequadamente treinados em material jurídico, o que denota que o enviesamento é esperado, ainda mais para modelos de IA que aparentemente têm como uma de suas funcionalidades primárias sustentar conversas com humanos (e nisso os modelos são bons).

Contudo, mesmo com os limites verificados, a capacidade de geração de texto coeso em termos linguísticos é evidente, motivo pelo qual a tecnologia de Processamento de Linguagem Natural presente no *GPT-3.5* e no *GPT-4* (e em outros sistemas congêneres que estão sendo desenvolvidos neste momento¹⁷⁵⁷) tem um potencial muitíssimo grande para aplicação na edição de textos jurídicos (resumo de argumentos, elaboração de relatórios, redação de tópicos específicos de documentos jurídicos e muito do que já foi descrito nesta obra), desde que haja um sistema auxiliar computacional (que pode ter uma IA própria treinada em Direito) que se aproveite das capacidades linguísticas do *GPT-3.5* e do *GPT-4* (ou novos modelos de IA que venham a ser desenvolvidos com escopo semelhante) e haja, também, o **efetivo e ininterrupto controle humano** sobre todos os textos produzidos com o auxílio da Inteligência Artificial (vide o subcapítulo 1.4), a fim de que seja formado o **duplo canal de correção** na via comunicacional humano-máquina (subcapítulo 1.6).

¹⁷⁵⁵ Graças ao acesso de desenvolvedor obtido em razão da presente pesquisa (OPENAI. **E-mail de OpenAI para Sérgio Rodrigo de Pádua com o título: “Welcome to the private beta for our API!”**, 2021. Acesso em 21 dez. 2022).

¹⁷⁵⁶ Por outro lado, considerada a experiência deste autor deve ser reconhecido que a *OpenAI* tem empreendido esforços de melhoria do *GPT-3.5* e do *GPT-4*, o que se revela numa maior segurança em relação a temas sensíveis ou linguagem imprópria que eventualmente apareciam nos dias iniciais de funcionamento do *GPT-3* e que aparentemente estão mais controlados nos dias atuais.

¹⁷⁵⁷ Para fins de registro histórico acerca do momento em que a pesquisa se refere, este trecho foi escrito em março de 2023.

Por exemplo, como saber se a IA utilizou o termo correto na elaboração de texto em Direito Constitucional?¹⁷⁵⁸ Ora, a curadoria dos resultados da IA deve ser reservada aos humanos que conheçam Direito Constitucional (e sua carga axiológica e filosófica), com o nível de expertise suficiente para adequar o texto ou propor outro que melhor corresponda à situação. Um problema (e que não está exatamente no escopo desta obra) seria garantir que um legítimo conhecedor do Direito esteja no controle da IA de maneira proativa e atenta, haja vista que um mero apertador de botões (ainda que tenha formação básica em Direito¹⁷⁵⁹) pode ser tentado a abandonar o canal humano da pretensão de correção em busca da solução mais fácil (a inércia típica do viés do *status quo* já descrita no subcapítulo 1.4). Educação jurídica (de qualidade) e tecnológica na formação de continuada de magistrados e servidores do Poder Judiciário é parte da resposta. Para além disso, as soluções de IA judicial baseadas em Processamento de Linguagem Natural (PLN) já fazem parte da realidade do Judiciário brasileiro (vide os dezenove projetos relatados pela Fundação Getúlio Vargas em 2022) e têm grande potencial de desenvolvimento, o que realça, ainda mais, a necessidade da garantia do controle humano voltado ao estabelecimento do duplo canal de correção.

¹⁷⁵⁸ Aqui o uso do Direito Constitucional é apenas um exemplo, pois o problema pode se repetir em qualquer ramo do Direito.

¹⁷⁵⁹ Infelizmente, na realidade brasileira, um diploma de Direito não é garantia que o portador conheça o Direito.

CONCLUSÃO

O juiz ciborgue foi investigado em ambas suas facetas, tanto o lado humano quanto a Inteligência Artificial foram esquadrihados e medidos, o que possibilita afastá-lo da figura de um julgador sobre-humano ou de um semideus. Assim, o juiz ciborgue atua consciente dos limites e das possibilidades do entrelaçamento Juiz-IA.

Sendo assim, a partir dos temas problematizados no texto, apresentam-se os seguintes resultados acerca dos elementos de legitimidade jurídico-constitucional do auxílio voltado à justificação da decisão judicial por meio de sistemas de Inteligência Artificial:

a) os juízes humanos decidem “racionalmente” por meio de heurísticas, maquinas também o fazem, o que indica espaço para o desenvolvimento de sistemas de IA voltados para auxílio à decisão judicial;

b) são necessários controle e supervisão humanos do sistema de IA judicial, a serem realizados pela pessoa do julgador, por sua equipe técnica e pelo Conselho Nacional de Justiça;

c) deve existir o respeito à pluralidade epistêmica voltada à pessoa do julgador, que pode utilizar a teoria do Direito (desde que democraticamente justificável) que lhe pareça a adequada segundo o seu posicionamento jurídico, uma vez que a teoria do Direito adotada pelo julgador pode influenciar no resultado prático de um julgamento;

d) ocorre a configuração de uma espécie de juiz ciborgue, nascido da interação que resulta do pensamento humano-máquina;

e) há possibilidade de maior objetividade dos julgamentos auxiliados por IA;

f) todos os riscos de utilização da IA judicial devem ser mapeados, tratados, monitorados de forma contínua e prevenidos;

g) estima-se que será cada vez mais necessária a figura do assessor-desenvolvedor (profissional com conhecimentos de Direito, Ciência da Computação e Ciência da Dados) para auxiliar o julgador/intérprete na implementação de modelos de IA no processo de decisão;

h) verifica-se a impossibilidade de centralização decisória mediante sistema computacional gerido exclusivamente pelo Tribunal, pela Corregedoria ou pelo Conselho Nacional de Justiça, em respeito à independência do julgador;

i) o Direito é também regido pelos processos informacionais (ainda que parte da teoria jurídica não trate sobre o tema), o que demanda o reconhecimento de sua mutabilidade inerente à entropia jurídica;

j) a partir da entropia do Direito e da dificuldade de sua representação estática (pelo julgador humano e pelos sistemas de IA judicial), há a necessidade de um duplo canal de correção, de maneira a demarcar vieses (e buscar eliminá-los) e a reduzir o ruído existente no pensamento humano ou nas soluções decisórias sugeridas por meio de IA, a fim de que o julgador humano possa filtrar as decisões propostas pela IA e esta possa fazer o mesmo em relação às decisões humanas (cotejo que é inerente ao processo argumentativo e à pretensão de correção);

k) a explicabilidade ocorre no limite do tecnicamente possível e se volta para a compreensão humana dos motivos de uma decisão tomada pela IA judicial, já a interpretabilidade pode ser limitada pelos aspectos inerentes à Ciência da Computação, uma vez que não se volta à compressão do cidadão comum;

l) o direito à explicabilidade das decisões tomadas com o auxílio de IA pode ser reconhecido como um direito fundamental por atribuição, o qual tem papel basilar ao possibilitar os ciclos do duplo canal de correção;

m) a IA pode funcionar em mais de uma camada decisória, na medida em que pode ser buscado um teste de juridicidade da solução implementada por IA, através de sistema sobreposto tipicamente voltado à melhora do grau de explicabilidade;

n) o grau de apoio da IA à decisão judicial pode variar (grau de complexidade e gestão de uso da IA);

o) frente a isso, a partir de agora o Direito deve traduzir o paradigma do *due process of law* para a era da IA;

p) a argumentação jurídica é um modelo que pode ser utilizado para garantir a legitimidade do uso da IA nas decisões judiciais, o que é pensado através dos modelos de justificação interna e justificação externa através de um procedimento de correção relativa;

r) o Direito pode ser parcialmente representado em sistemas de Inteligência Artificial judicial, voltados à argumentação jurídica necessária para a formação do duplo canal de correção, representação esta que leva em conta a justificação interna e a justificação externa típicas do discurso racional;

s) a participação do julgador/intérprete humano e de sua equipe técnica (o que inclui o assessor-desenvolvedor) no processo de desenvolvimento dos sistemas de IA judicial é parte integrante da pretensão de correção, uma vez que o processo de desenvolvimento típico da Ciência de Dados aplicada ao Direito exige que haja efetiva contribuição dos juristas e operadores do Direito, de modo que sistemas de IA argumentativa e juridicamente mais eficientes possam ser implementados.

Sendo assim, desde que implementado o duplo canal de correção com todos os seus elementos definidores, a IA judicial possibilitará a constituição do juiz ciborgue, o qual se volta à correta aplicação do Direito. Logo, neste momento da conclusão, há a oportunidade para a delimitação final a respeito do duplo canal de correção.

Conforme exposto no subcapítulo 1.6, uma vez reconhecida a típica relatividade dos discursos reais, com seus limites que foram explorados no decorrer desta obra, o entrelaçamento Juiz-IA resulta em um sistema híbrido de pensamento humano-máquina, haja vista a via comunicacional que é inerente ao duplo canal de correção, sendo este necessário ao estabelecimento do modelo de juiz ciborgue. Assim, mais uma vez, veja-se o diagrama do duplo canal de correção:

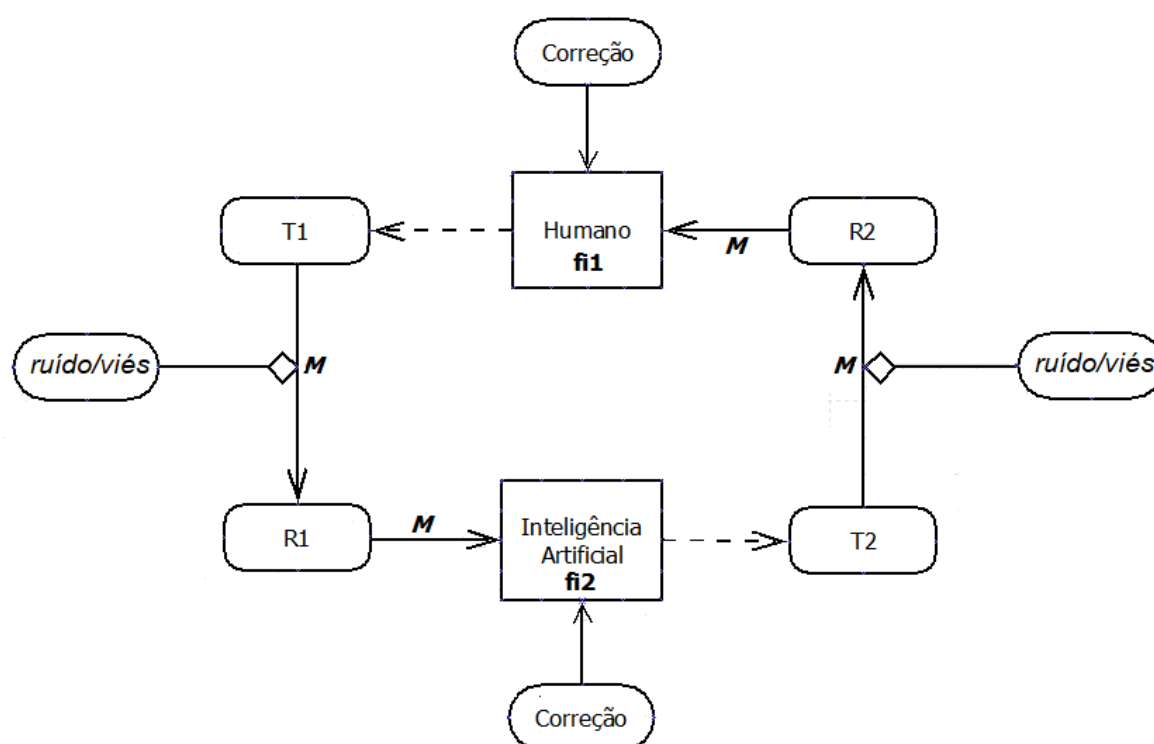


Fig. 15 – Diagrama Esquemático do Duplo Canal de Correção.

Os detalhes de funcionamento do duplo canal de correção foram expostos no subcapítulo 1.6, motivo pelo qual é desnecessária uma repetição integral neste momento. Todavia, a partir das considerações já realizadas nesta conclusão, devem ser definidos os elementos básicos de funcionamento do duplo canal de correção.

Portanto, a descrição básica do funcionamento do processo argumentativo-comunicacional interno que ocorre no duplo canal de correção engloba os seguintes elementos:

1. No funcionamento interno do duplo canal de correção, o destinatário da mensagem (Humano em **fi1** ou IA em **fi2**) é também corretor (tem função de correção), haja vista o aspecto de mútua colaboração para a redução de erros, ruídos e vieses;

2. Humano e IA (Inteligência Artificial) podem buscar qualquer argumento juridicamente admissível para o estímulo dos ciclos do duplo canal de correção, aspecto que leva em consideração a legítima possibilidade de representação do Direito, através de proposições normativas que englobam aspectos da Teoria do Direito (processo de individualização da norma jurídica e argumentação jurídica) e da Ciência da Computação (conceitos jurídicos intermediados e fatores);

3. Ao menos uma interação humano-máquina deve ser realizada através do duplo canal de correção, sendo que a correção não é uma garantia de única resposta possível, pois se trata de uma correção procedimental relativa (engloba correção normativa e correção moral) que se sujeita a um grau de incerteza advinda da entropia jurídica;

4. Os ciclos de interação humano-máquina devem ser realizados até que haja uma justificativa humana que supere argumentativamente as soluções propostas pela IA, o que inclui a explicabilidade da decisão e dos motivos pelos quais as propostas da Inteligência Artificial foram aceitas ou rejeitadas, eis que se demonstra a configuração do direito fundamental à explicabilidade como uma premissa ao funcionamento do duplo canal de correção;

5. Os ciclos de correção podem ser realizados até que o intérprete humano entenda pela juridicidade da proposição obtida, pois deve ser mantido o controle humano sobre as decisões judiciais auxiliadas por sistemas de Inteligência Artificial, a fim de se possibilitar a legitimidade jurídica do uso da IA em decisões judiciais;

6. As sugestões de aplicação do Direito em cada ciclo, humanas e computacionais, devem ser registradas e explicáveis, uma vez que a pretensão de

correção só é comprovável quando há o pleno respeito ao direito fundamental à explicabilidade;

7. No limite, o Humano decide quando se encerram os ciclos de correção, desde que estritamente respeitadas as demais regras de funcionamento do duplo canal de correção, haja vista os necessários controle e supervisão humanos;

8. A partir das considerações existentes no Capítulo 5, o desenvolvimento da IA voltada ao duplo canal de correção deve contar com a participação dos julgadores/intérpretes e de suas equipes técnicas (inclusive dos assessores-desenvolvedores), tendo em vista o paradigma da Ciência de Dados que, juntamente com o Direito, é aplicável em relação às técnicas e aos processos de construção de sistemas de Inteligência Artificial judicial.

Diante da presente tese, o desafio que se coloca para os futuros pesquisadores do tema da Inteligência Artificial judicial é o desenvolvimento de sistemas que consigam aproveitar e otimizar o duplo canal de correção, de maneira a levar a pretensão de correção aos seus limites.

Logo, o juiz ciborgue mantém sua humanidade, mas busca repelir seus ímpetos irracionais pelos meios que a tecnologia lhe oferece. Sobretudo, o juiz ciborgue se trata de um mortal, que ainda sente e sangra e que é avesso à alcunha de portador último da “verdade”, do “bem”, da “felicidade”, da “bondade” e da “suprema racionalidade”. O juiz ciborgue é um juiz que chora e que ri. Ele é o juiz que vive com a sua irrevogável humanidade.

REFERÊNCIAS

AFONSO DA SILVA, Virgílio. **Direitos Fundamentais**: conteúdo essencial, restrições e eficácia. São Paulo: Malheiros, 2009.

_____. O proporcional e o razoável. **Revista dos Tribunais**, n. 798, p. 23-50, 2002.

ALCHOURRÓN, Carlos E.; BULYGIN, Eugenio. **Normative Systems**. New York: Springer-Verlag, 1971.

ALEXY, Robert. **Conceito e Validade do Direito**. Tradução: Gercélia Batista de Oliveira Mendes. São Paulo: WMF Martins Fontes, 2009.

_____. **Direito, Razão, Discurso**: Estudos para a filosofia do direito. Tradução: Luís Afonso Heck. 2ª ed. Porto Alegre: Livraria do Advogado, 2015.

_____. **Princípios Formais**. Rio de Janeiro: Forense, 2018. E-book Kindle.

_____. Sistema Jurídico, Principios Jurídicos y Razón Práctica. **Doxa**, Madrid, n. 5, 1988.

_____. Sobre Duas Justaposições: Conceito e Natureza, Direito e Filosofia. Alguns Comentários sobre “Pode Haver Uma Teoria do Direito?” de Joseph Raz. In: RAZ, Joseph. **Uma Discussão Sobre Teoria do Direito**. Tradução: Sheila Stolz. São Paulo: Marcial Pons, 2013.

_____. **Teoria da Argumentação Jurídica**: A Teoria do Discurso Racional como Teoria da Justificação Jurídica. 6.ª ed. Tradução: Zilda Hutchinson Schild Silva. Revisão Técnica e Apresentação: Cláudia Toledo. Rio de Janeiro: Forense, 2021.

_____. **Teoria de Los Derechos Fundamentales**. Madri: Centro de Estudios Constitucionales, 1997.

_____. **Teoria Discursiva do Direito**. 3ª ed. Tradução: Alexandre Travessoni Gomes Trivisonno. Rio de Janeiro: Forense, 2018. E-book Kindle.

_____. **Teoria dos Direitos Fundamentais**. 2ª ed. Tradução: Virgílio Afonso da Silva. São Paulo: Malheiros, 2011.

_____. The Absolute and the Relative Dimension of Constitutional Rights. **Oxford Journal of Legal Studies**, Oxford, p. 1-17, 2016. DOI: 10.1093/ojls/gqw013.

ALMEIDA, Adriano; CARVALHO, Felipe; MENINO, Felipe; NETO, Helvécio. **Introdução ao Machine Learning**. [2017]. Disponível em: dataat.github.io/introducao-ao-machine-learning/introducao.html. Acesso em 31 jan. 2023.

ALVES, Isabella Fonseca; ALMEIDA, Priscila Brandão de. Direito 4.0: uma análise sobre inteligência artificial processo e tendências de mercado. In: ALVES, Isabella Fonseca (org.). **Inteligência Artificial e Processo**. Belo Horizonte: D’Plácido, 2020.

ANGWIN, Julia et al. **Machine Bias**. Disponível em: www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. Acesso em 28 jun. 2022.

ARAÚJO, Erik da Silva e; SIMIONI, Rafael Larazzotto. A Inteligência Artificial e Ponderação: o Procedimentalismo de Robert Alexy na Forma de Algoritmo. **Revista Jurídica Eletrônica da Universidade Federal do Piauí**, v. 8, n. 1, jan/jun 2021.

ASIMOV, Isaac. **Eu, Robô**. Tradução: Aline Storto Pereira. São Paulo: Aleph, 2015. E-book Kindle.

ASHLEY, Kevin D. **Artificial Intelligence and Legal Analytics**: New tools for law practice in digital age. Cambridge: Cambridge University Press, 2017.

_____. Case-based reasoning and its implications for legal expert systems. **Artificial Intelligence and Law**, Dordrecht, v. 1, p. 113-208, 1992, DOI 10.1007/BF00114920.

_____. Evaluating the Uses of Values in a Model of Legal Reasoning. In: ATKINSON, Katie; PRAKKEN, Henry; WYNER, Adam. **From Knowledge Representation to Argumentation in AI, Law and Policy Making**: A Festschrift in Honour of Trevor Bench-Capon on Occasion of his 60th Birthday. College Publications: London, 2013.

ATIENZA, Manuel. **As Razões do Direito**: Teoria da Argumentação Jurídica. Tradução: Maria Cristina Guimarães Cupertino. 2ª ed. Rio de Janeiro: Forense Universitária, 2016.

ÁVILA, Humberto. **Teoria dos Princípios**: da definição à aplicação dos princípios jurídicos. 5ª ed. São Paulo: Malheiros, 2006.

AWS. **Adicione facilmente inteligência às suas aplicações**. [2022]. Disponível em: <https://aws.amazon.com/pt/machine-learning/ai-services>. Acesso em 15 nov. 2022.

BARBOZA, Estefânia Maria de Queiroz. **Stare decisis, integridade e segurança jurídica**: reflexões críticas a partir da aproximação dos sistemas de common law e civil law. 2010. Tese (Doutorado em Direito). Pontifícia Universidade Católica do Paraná (PUC-PR), Curitiba, 2010.

BARROSO, Luís Roberto. **Interpretação e Aplicação da Constituição**. 3ª ed. São Paulo: Saraiva, 1999.

_____. Os Princípios da Razoabilidade e da Proporcionalidade no Direito Constitucional. **BDA – Boletim de Direito Administrativo**, p. 156-166, março/1997.

BAUDRILLARD, Jean. **Simulacros e Simulação**. Lisboa: Relógio d'Água, 1991.

BENCH-CAPON, Trevor. **Knowledge-Based Systems and Legal Applications**. London: Academic Press, 2015.

BENCH-CAPON, Trevor et al. A history of AI and Law in 50 papers: 25 years of the international conference on AI and Law. **Artificial Intelligence and Law**, Dordrecht, v. 20, p. 215-319, 2012. DOI: 10.1007/s10506-012-9131-x.

BENCH-CAPON, Trevor; SARTOR, Giovanni. A Model of Legal Reasoning with Cases Incorporating Theories and Values. **Artificial Intelligence**, v. 150, n. 1-2, p. 97-143, 2003.

BELLAMY, Matthew. Algorithm. In: MUSE. **Simulation Theory**. Londres: Warner Bros. Records e Helium-3, 2018. Disponível em: <https://youtu.be/wJcVLxRtDms>. Acesso em 01 jun. 2022.

BERG, Lyudmila N. System of Legal Impact from Social Entropy to Legal Development. **Revista Quid**, v. 2, Special Issue, p. 133-137, 2018.

BITTAR, Eduardo Carlos Bianca; ALMEIDA, Guilherme de Assis de. **Curso de Filosofia do Direito**. 11ª ed. São Paulo: Atlas, 2015.

BOBBIO, Norberto. **Teoria Geral do Direito**. Tradução: Denise Agostinetti. 2ª ed. São Paulo: Martins Fontes, 2008.

BOEING, Daniel Henrique Arruda; DA ROSA, Alexandre Moraes. **Ensinando um Robô a Julgar**: Pragmática, Discricionariedade, Heurísticas e Vieses no Uso de Aprendizado de Máquina no Judiciário. Florianópolis: EMais Academia, 2020.

BOFF, Salete Oro; FORTES, Vinícius Borges; FREITAS, Cinthia Obladen de Almendra. **Proteção de Dados e Privacidade**: Do Direito às Novas Tecnologias na Sociedade da Informação. Rio de Janeiro: Lumen Juris, 2018.

BONAT, Debora; HARTMANN PEIXOTO, Fabiano. **Racionalidade no Direito**: Inteligência Artificial e Precedentes. Curitiba: Alteridade, 2020.

BOSTROM, Nick. **Superinteligência**: Caminhos, perigos, estratégias. Tradução: Aurélio Antônio Monteiro, Clemente Gentil Penna, Fabiano Geremias Monteiro e Patrícia Ramos Geremias. Rio de Janeiro: Darkside Books, 2014. E-book Kindle.

BRASIL. **Decreto 678, de 06 de novembro de 1992**. Convenção Americana de Direitos Humanos. Disponível em: http://www.planalto.gov.br/ccivil_03/decreto/d0678.htm. Acesso em 28 dez. 2022.

_____. **Emenda Constitucional nº 45, de 30 de dezembro de 2004**. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/emendas/emc/emc45.htm. Acesso em 12 jun. 2022.

_____. **Lei 7.713, de 22 de março de 1988**. Altera a legislação do imposto de renda e dá outras providências. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/l7713.htm. Acesso em: 30 nov. 2022.

_____. **Lei 10.406, de 10 de janeiro de 2002**. Código Civil. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/2002/l10406compilada.htm. Acesso em: 27 nov. 2022.

_____. **Lei 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: 11 jun. 2022.

_____. **Lei nº 13.655, de 25 de abril de 2018.** Disponível em: https://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2018/Lei/L13655.htm Acesso em: 11 jun. 2022.

_____. **Lei 13.709, de 14 de agosto de 2018.** Lei Geral de Proteção de Dados. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 11 jun. 2022.

_____. Lei Complementar nº 35, de 14 de março de 1979. Dispõe sobre a Lei Orgânica da Magistratura Nacional. **Diário Oficial da União:** Brasília, DF, 14 mar. 1979.

_____. Senado Federal. **Projeto de Lei nº 21, de 2020.** Estabelece fundamentos, princípios e diretrizes para o desenvolvimento e a aplicação da inteligência artificial no Brasil; e dá outras providências. Brasília. Senado Federal, 03 fev. 2022. Disponível em: legis.senado.leg.br/sdleg-getter/documento?dm=9063365&ts=1656528542410&disposition=inline. Acesso em: 17 jul. 2022.

_____. Superior Tribunal de Justiça. Resolução STJ/GP nº 10, de 6 de outubro de 2015. Regulamenta o processo judicial eletrônico no Superior Tribunal de Justiça. **Diário de Justiça Eletrônico do STJ:** Brasília, DF, 07 out. 2015.

_____. Superior Tribunal de Justiça. **STJ entra na era da inteligência artificial.** [2018]. Disponível em: http://www.stj.jus.br/sites/portaltj/Paginas/Comunicacao/Noticias-antigas/2018/2018-06-14_06-51_STJ-entra-na-era-da-inteligencia-artificial.aspx. Acesso em: 01 set. 2019.

_____. Superior Tribunal de Justiça. **Súmula 486 do STJ.** Disponível em: <https://www.stj.jus.br/publicacaoainstitucional/index.php/sumstj/article/download/5174/5299>. Acesso em: 20 jul. 2022.

_____. Superior Tribunal de Justiça (2. Turma), **Agravo Interno no Recurso Especial 1786179/SP.** Não cabimento de Recurso Especial em face de acórdão de ação rescisória cujo fundamento de julgamento se baseou na aplicação de dispositivo constitucional. Brasília, DF: Superior Tribunal de Justiça, [2019]. Recorrente: Zahir Dornaika e outros. Recorrido: Fazenda do Estado de São Paulo. Relator: Min. Herman Benjamin, 06 de junho de 2019. Disponível em: https://scon.stj.jus.br/SCON/GetInteiroTeorDoAcordao?num_registro=201803254544&dt_publicacao=18/06/2019. Acesso em: 14 nov. 2021.

_____. Superior Tribunal de Justiça (3. Turma). **Agravo Regimental nos Embargos de Declaração no Agravo em Recurso Especial 418.785/RJ.** Aplicação do Código de Defesa do Consumidor aos casos de responsabilidade civil decorrente de transporte aéreo internacional em detrimento das normas da Convenção de Haia e Convenção de Montreal. Brasília, DF: Superior Tribunal de Justiça, [2016]. Relator: Min. João Otávio de Noronha, 17 de maio de 2016. Disponível em:

https://scon.stj.jus.br/SCON/GetInteiroTeorDoAcordao?num_registro=201303584762&dt_publicacao=23/05/2016. Acesso em: 09 jul. 2022.

_____. Superior Tribunal de Justiça (4. Turma). **Agravo Interno no Agravo em Recurso Especial nº 1.718.713/GO**. Acórdão de decisão que negou provimento ao pedido de revisão de valor de indenização por dano moral. Brasília, DF: Superior Tribunal de Justiça, [2020]. Recorrente: Danglacia Teofila Constantino. Recorridos: Bradesco Auto e outros. Relator: Ministro Marco Buzzi, 26 de outubro de 2020. Disponível em: https://scon.stj.jus.br/SCON/GetInteiroTeorDoAcordao?num_registro=202001502389&dt_publicacao=12/11/2020. Acesso em: 10 ago. 2021.

_____. Superior Tribunal de Justiça (6. Turma). **Recurso Especial 1177612/SP. Habeas corpus** concedido de ofício em razão da não configuração da tipicidade subjetiva para o crime de falsificação de documentos públicos (decisões judiciais). Brasília, DF: Superior Tribunal de Justiça, [2011]. Relator: Min. Og Fernandes, 01 de setembro de 2011. Disponível em: https://scon.stj.jus.br/SCON/GetInteiroTeorDoAcordao?num_registro=201000156359&dt_publicacao=17/10/2011. Acesso em: 12 jan. 2023.

_____. Supremo Tribunal Federal. **Inteligência artificial vai agilizar a tramitação de processos no STF**. [2018]. Disponível em: <http://www.stf.jus.br/portal/cms/verNoticiaDetalhe.asp?idConteudo=380038>. Acesso em: 01 set. 2019.

_____. Supremo Tribunal Federal. Resolução nº 427, de 20 de abril de 2010. Regulamenta o processo eletrônico no âmbito do Supremo Tribunal Federal. **Diário de Justiça Eletrônico do STF**: Brasília, DF, nº 72, p. 1, 26 abr. 2010.

_____. Supremo Tribunal Federal. **Pesquisa de Jurisprudência**. [2021]. Disponível em: <https://jurisprudencia.stf.jus.br>. Acesso em: 31 jul. 2021.

_____. Supremo Tribunal Federal (1. Turma), **Recurso Ordinário em Habeas Corpus 165393 AgR/PE**. Ato alheio à jurisdição do Supremo Tribunal Federal por meio de Recurso Ordinário, bem como incabível a análise do *habeas corpus* (rol taxativo do art. 102, I, “d” e “i”, da Constituição Federal). Brasília, DF: Supremo Tribunal Federal, [2019]. Recorrente: Eudes Teixeira de Carvalho Júnior. Recorrido: Ministério Público Federal. Relator: Min. Luiz Fux, 15 de março de 2019. Disponível em: <https://jurisprudencia.stf.jus.br/pages/search/sjur400729/false>. Acesso em: 10 dez. 2021.

_____. Supremo Tribunal Federal (2. Turma). **Agravo de Instrumento 548203 ED/AL**. Nulidade absoluta do processo por violação ao princípio do juiz natural. Brasília, DF: Supremo Tribunal Federal, [2008]. Recorrente: Estado de Alagoas. Recorrido: Vanildo de Magalhães Maurício. Relator: Min. Cezar Peluso, 12 de fevereiro de 2008. Disponível em: <https://jurisprudencia.stf.jus.br/pages/search/sjur88959/false>. Acesso em: 10 dez. 2021.

_____. Supremo Tribunal Federal (2. Turma). **Recurso Ordinário em Habeas Corpus 144615 AgR/PR**. Inobservância do sistema acusatório no processo em que o juiz tem postura ativa na instrução processual. Brasília, DF: Supremo Tribunal Federal, [2020]. Recorrente: Paulo Roberto Krug. Recorrido: Ministério Público Federal. Relator: Min. Edson Fachin. Relator para acórdão: Min. Gilmar Mendes, 25 de agosto

de 2020. Disponível em: <https://jurisprudencia.stf.jus.br/pages/search/sjur435485/false>. Acesso em: 10 dez. 2021.

_____. Supremo Tribunal Federal (Plenário). **Ação Direta de Inconstitucionalidade 4357**. Inconstitucionalidade parcial da Emenda Constitucional 62/2009, em razão da ofensa ao direito à propriedade. Brasília, DF: Supremo Tribunal Federal, [2013]. Relator: Min. Carlos Ayres Britto. Relator para acórdão: Min. Luiz Fux, 14 de março de 2013. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=6812428>. Acesso em: 21 abr. 2021.

_____. Supremo Tribunal Federal (Plenário). **Ação Direta de Inconstitucionalidade 4425**. Inconstitucionalidade parcial da Emenda Constitucional 62/2009, em razão da ofensa ao direito à propriedade, à isonomia, à efetividade da jurisdição, à coisa julgada material, à separação dos poderes, fixação de critérios atualização monetária para dívidas públicas e de critérios para pagamento de precatórios. Brasília, DF: Supremo Tribunal Federal, [2013]. Relator: Min. Luiz Fux, 14 de março de 2013. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=5067184>. Acesso em: 21 abr. 2021.

_____. Supremo Tribunal Federal (Plenário). **Ação Direta de Inconstitucionalidade 4451**. Inconstitucionalidade do art. 45, II e III, e §4º e §5º, da Lei 9.504/1997, Lei de Eleições, em razão da ofensa à liberdade de manifestação do pensamento, à liberdade de comunicação e expressão artística e ao livre exercício profissional de humoristas. Brasília, DF: Supremo Tribunal Federal, [2018]. Relator: Min. Alexandre de Moraes, 21 de junho de 2018. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=749287337>. Acesso em: 21 abr. 2021.

_____. Supremo Tribunal Federal (Plenário). **Agravo de Instrumento 791.292 QO-RG (Tema 339 de Repercussão Geral)**. O dever de fundamentação das decisões judiciais não pressupõe o enfrentamento de todas as teses alegadas pelas partes e nem mesmo a correção dos fundamentos utilizados. Brasília, DF: Supremo Tribunal Federal, [2010]. Relator: Min. Gilmar Mendes, 23 de junho de 2010. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=AC&docID=613496>. Acesso em: 11 jun. 2022.

_____. Supremo Tribunal Federal (Plenário). **Agravo Regimental no Recurso Extraordinário 820.433**. Constitucionalidade de desersão de recurso reconhecida pelo Superior Tribunal de Justiça mediante a aplicação do art. 93, XIV, da Constituição Federal. Brasília, DF: Supremo Tribunal Federal, [2016]. Relator: Min. Dias Toffoli, 17 de março de 2016. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=11043891>. Acesso em: 12 jan. 2023.

_____. Supremo Tribunal Federal (Plenário). **Arguição de Descumprimento de Preceito Fundamental 130**. Não recepção da Lei de Imprensa, Lei 5.250/1967, pela Constituição Federal de 1998, em razão da incompatibilidade com a liberdade de opinião, a liberdade de informação e comunicação e a liberdade de imprensa. Brasília, DF: Supremo Tribunal Federal, [2009]. Relator: Min. Carlos Ayres Britto, 30 abril de

2009. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=AC&docID=605411>. Acesso em: 21 abr. 2021.

_____. Supremo Tribunal Federal (Plenário). **Medida Cautelar em Ação Direta de Inconstitucionalidade 6387**. Defere medida cautelar para afastar os efeitos da Medida Provisória nº 954/2020 em razão da ofensa ao direito fundamental à autodeterminação informativa. Brasília, DF: Supremo Tribunal Federal, [2020]. Relator: Min. Gilmar Mendes, 07 de maio de 2020. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=754357629>. Acesso em: 28 dez. 2022.

_____. Supremo Tribunal Federal (Plenário). **Questão de Ordem na Ação Direta de Inconstitucionalidade 4425**. Modulação dos efeitos da declaração de inconstitucionalidade parcial da Emenda Constitucional 62/2009. Relator: Min. Luiz Fux: 25 de março de 2015. Brasília, DF: Supremo Tribunal Federal, [2015]. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=9016259>. Acesso em: 21 abr. 2021.

_____. Supremo Tribunal Federal (Plenário). **Questão de Ordem no Inquérito 2503/SP**. Nulidade de sentença pela ausência de imparcialidade do juiz. Brasília, DF: Supremo Tribunal Federal, [2008]. Querelante: José Ursílio de Souza e Silva. Investigado: José Abelardo Guimarães Camarinha. Relator: Min. Menezes Direito, 13 de março de 2008. Disponível em: <https://jurisprudencia.stf.jus.br/pages/search/sjur3049/false>. Acesso em: 10 dez. 2021.

_____. Supremo Tribunal Federal (Plenário). **Recurso Extraordinário 4357**. Hierarquia Supralegal dos tratados e convenções sobre Direitos Humanos. Brasília, DF: Supremo Tribunal Federal, [2009]. Relator: Min. Carlos Ayres Britto. Relator para acórdão: Min. Gilmar Mendes, 03 de dezembro de 2008. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=AC&docID=595406>. Acesso em: 30 nov. 2022.

_____. Supremo Tribunal Federal (Plenário). **Recurso Extraordinário 636.331**. As Convenções Haia e de Montreal têm prevalência sobre o Código de Defesa do Consumidor e são aplicáveis aos casos de responsabilidade civil em contratos de transporte aéreo internacional. Brasília, DF: Supremo Tribunal Federal, [2017]. Relator: Min. Gilmar Mendes, 25 de maio de 2017. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=14028416>. Acesso em: 09 jul. 2022.

_____. Supremo Tribunal Federal (Plenário). **Recurso Extraordinário 1.394.401**. Em matéria de danos extrapatrimoniais relacionados ao transporte aéreo internacional, o Código de Defesa do Consumidor tem prevalência sobre as Convenções Varsóvia e de Montreal. Brasília, DF: Supremo Tribunal Federal, [2022]. Relator: Min. Rosa Weber, 15 de dezembro de 2022. Disponível em: <https://portal.stf.jus.br/processos/downloadPeca.asp?id=15356316433&ext=.pdf>. Acesso em: 04 mar. 2023

_____. Supremo Tribunal Federal (Plenário). **Recurso Extraordinário com Agravo 654432 (Tema 541 de Repercussão Geral)**. Fixa o Tema 451 de repercussão geral, mediante a interpretação teleológica dos art. 9º, §1º, art. 37, VII, e

art. 144, da Constituição Federal, a fim de estabelecer proibição absoluta ao exercício do direito de greve a todos os servidores públicos integrantes das carreiras de segurança pública. Brasília, DF: Supremo Tribunal Federal, [2017]. Relator: Min. Edson Fachin. Relator para acórdão: Min. Alexandre de Moraes, 05 de abril de 2017. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=14980135>. Acesso em: 22 abr. 2021.

_____. Supremo Tribunal Federal (Plenário). Voto-vista do Ministro Sepúlveda Pertence, na Ação Direta de Inconstitucionalidade nº 2. Possibilidade de conhecimento de Ação Direta de Inconstitucionalidade com base da teoria da inconstitucionalidade superveniente. Brasília, DF: Supremo Tribunal Federal, [2020]. Autores: Federação Nacional dos Estabelecimentos de Ensino e Presidente da República. Relator: Ministro Paulo Brossard, 06 de fevereiro de 1992. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=AC&docID=266151>. Acesso em: 10 ago. 2021.

_____. Tribunal Regional do Trabalho da 1ª Região (10. Turma). **Recurso Ordinário 0010600-02.2015.5.01.0261/RJ**. Interpretação *in dubio pro operario* de cláusula normativa. Rio de Janeiro, RJ: Tribunal Regional do Trabalho da 1ª Região, [2017]. Relator: Des. Flavio Ernesto Rodrigues Silva, 26 de abril de 2017. Disponível em: <https://bibliotecadigital.trt1.jus.br/jspui/handle/1001/1343580?mode=full>. Acesso em: 12 dez. 2021.

_____. Tribunal Regional Federal da 4ª Região. **Eproc: processo eletrônico da Justiça Federal da 4ª Região completa cinco anos**. [2014]. Disponível em: https://www.trf4.jus.br/trf4/controlador.php?acao=noticia_visualizar&id_noticia=10500. Acesso em 01 mai. 2020.

_____. Tribunal Regional Federal da 4ª Região. Resolução nº 13, de 11 de março de 2004. Implanta e estabelece normas para o funcionamento do Processo Eletrônico nos Juizados Especiais Federais no âmbito da Justiça Federal da 4ª Região. **Boletim Interno Eletrônico**: Porto Alegre, RS, n.º 174-06, 20 abr. 2004.

_____. Tribunal Regional Federal da 4ª Região. **Resolução nº 64, de 11 de novembro de 2009**. Implanta o Processo Eletrônico no âmbito da Justiça Federal de 1º e 2º Graus da 4ª Região. Porto Alegre. TRF4, 11 nov. 2009. Disponível em: <https://www.trf4.jus.br/trf4/sup/res64-2009.pdf>. Acesso em: 02 fev. 2022.

_____. Tribunal Superior do Trabalho. **Histórico do PJe**. [2020]. Disponível em: <http://www.tst.jus.br/web/pje/historico>. Acesso em 01 mai. 2020.

BRETAS, Carla Panza; MATTOS, Kennedy Josue Greca de; SCHIER, Paulo Ricardo. Teoria do Direito e da Constituição no contexto do neoconstitucionalismo pós-positivista: análise quanto à racionalidade. **Revista da Faculdade de Direito da UFRGS**, Porto Alegre, v. 35, p. 115-130, vol. Esp., dez. 2016.

BROWN, Tom B. et al. **Language Models are Few-Shot Learners**. OpenAI, San Francisco, arXiv:2005.14165. Disponível em: <https://arxiv.org/pdf/2005.14165.pdf>. Acesso em: 05 abr. 2021.

BULYGIN, Eugenio. Alexy e o Argumento da Correção. In: BULYGIN, Eugenio; ALEXY, Robert. **Direito, Moralidade e Positivismo: o debate entre Robert Alexy e Eugenio Bulygin**. Curitiba: Juruá, 2017.

_____. **Lógica Deóntica, Normas y Propositiones Normativas**. Madri: Marcial Pons, 2018.

_____. Raz e a Teoria do Direito. Comentários Sobre “Pode Haver Uma Teoria do Direito?” de Joseph Raz. In: RAZ, Joseph. **Uma Discussão Sobre Teoria do Direito**. Tradução: Sheila Stolz. São Paulo: Marcial Pons, 2013.

CAMPOS, Anna Maria. Accountability: quando poderemos traduzi-la para o português? **Revista de Administração Pública**, Rio de Janeiro, v. 24, n. 2, p. 30-50, jun. 1990.

CAMPOS, Ricardo. Metamorfoses do Direito Global: sobre a interação entre direito, tempo e tecnologia. São Paulo: Contracorrente, 2022.

CARNEIRO, Claudio; LEMOS, Marcos Antonio Fernandes. O Movimento Pós-Positivista e a “Visão” Neoconstitucionalista da Ponderação: A Valoração de Princípios em Robert Alexy. **Revista de Direitos Fundamentais & Democracia**, Curitiba, v. 12, n. 12, jul./dez. 2012.

CASETEXT. **CARA A.I.** [2021]. Disponível em: <https://casetext.com/cara-ai>. Acesso em: 28 abr. 2021.

CAPPELLETTI, Mauro; GARTH, Bryant. **Acesso à Justiça**. Tradução: Ellen Gracie Northfleet. Porto Alegre: Sérgio Fabris, 1988.

CELLA, José Renato Gaziero. **Controle das Decisões Jurídicas pela Técnica do Autoprecedente**: Lógica Deôntica Paraconsistente Aplicada em Sistemas Especialistas Legais. Tese (Doutorado em Direito) – Universidade Federal de Santa Catarina, Florianópolis, p. 268, 2008.

CELLA, José Renato Gaziero; WOJCIECHOWSKI, Paola Bianchi. Inteligência Artificial nos Processos Judiciais Eletrônicos. In: FEITOSA, Raymundo Juliano Rego (Org.); et. al. **Direito e novas tecnologias**, Coleção Conpedi/Unicuritiba. Vol.12. São Paulo: Clássica Editora, 2014.

CHAKRABARTI, C.; CHAKRABARTY, Indranil. Boltzmann Entropy: Probability and Information. **Romanian Journal of Physics**, Bucharest, v. 52, n. 5–7, p. 559–564, 2007.

CHALKIDIS, Ilias; KAMPAS, Dimitrios. Deep Learning in Law: Early Adaptation and Legal Word Embeddings Trained On Large Corpora. **Artificial Intelligence and Law**, Dordrecht, v. 27, p. 171-198, jun. 2019. DOI: 10.1007/s10506-018-9238-9.

CHATGPT. **Resposta à pergunta: Você conhece Direito Constitucional do Brasil?** Disponível em: <https://chat.openai.com/chat>. Acesso em 20 dez. 2022.

_____. **Resposta à pergunta: Você pode explicar a teoria da simetria constitucional?** Disponível em: <https://chat.openai.com/chat>. Acesso em 20 dez. 2022.

_____. **Teoria da Justiça em poucas palavras.** Disponível em: <https://chat.openai.com/chat>. Acesso em 20 dez. 2022.

_____. **Theory of Justice in simple terms.** Disponível em: <https://chat.openai.com/chat>. Acesso em 20 dez. 2022.

CHOULDECHOVA, Alexandra. Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments. **Big Data**, v. 5, n. 2, p. 153-1163, 2017. DOI: 10.1089/big.2016.0047.

CHRISTIAN, Brian; GRIFFITHS, Tom. **Algoritmos para viver**. Tradução: Paulo Geiger. São Paulo: Companhia das Letras, 2016. E-book Kindle.

CINTRA, Antonio Carlos de Araújo; GRINOVER, Ada Pellegrini; DINAMARCO, Cândido Rangel. **Teoria Geral do Processo**. 19ª ed. São Paulo: Malheiros, 2003.

CLAY MATHEMATICS INSTITUTE. **P vs NP Problem**. [2020]. Disponível em: <https://www.claymath.org/millennium-problems/p-vs-np-problem>. Acesso em: 11 jan. 2022.

CLYNES, Manfred E.; KLINE, Nathan S. Cyborgs and Space. **Astronautics**, p. 29-33, Sep. 1960.

CONSELHO NACIONAL DE JUSTIÇA. **Área de Atuação do Projudi**. [2020]. Disponível em: <https://www.cnj.jus.br/area-de-atuacao-projudi>. Acesso em: 01 mai. 2020.

_____. **CNJ anuncia a criação de laboratório de inteligência artificial para o PJe**. [2018]. Disponível em: <https://www.cnj.jus.br/noticias/cnj/88132-cnj-anuncia-a-criacao-de-laboratorio-de-inteligencia-artificial-para-o-pje>. Acesso em: 31 ago. 2019.

_____. **Datajud: Base Nacional de Dados do Poder Judiciário**. [2022]. Disponível em: <https://www.cnj.jus.br/sistemas/datajud/>. Acesso em: 12 set. 2022.

_____. **Histórico do PJe**. [2020]. Disponível em: https://www.cjf.jus.br/observatorio/arq/historico_pje.pdf. Acesso em: 01 mai. 2020.

_____. **Módulo de Produtividade Mensal do Sistema de Estatísticas do Poder Judiciário**. [2022]. Disponível em: <https://www.cnj.jus.br/capg/mapa.php>. Acesso em: 11 jun. 2022.

_____. **Portaria nº 196, de 10 de outubro de 2022**. Institui regras de acesso das equipes técnicas ao DataJud. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3428>. Acesso em: 12 set. 2022.

_____. **Procedimento de Controle Administrativo 0000303- 58.2011.2.00.0000**. Inconstitucionalidade da completa delegação de atos jurisdicionais a Juízes Leigos. Brasília, DF: Conselho Nacional de Justiça, [2011]. Relator: Cons. Jefferson Luis

Kravchychyn, 01 de março de 2011. Disponível em: <https://www.cnj.jus.br/Infojurisl2/Jurisprudencia.seam;jsessionid=1463BB48E76D0A3726156EF32E558E60?jurisprudencialdJuris=42662&indiceListaJurisprudencia=6&firstResult=900&tipoPesquisa=BANCO>. Acesso em: 11 jun. 2022.

_____. Recomendação nº 51, de 24 de março de 2015. Recomenda a utilização dos Sistemas Bacenjud, Renajud e Infojud e dá outras providências. **Diário da Justiça**: Brasília, DF, n.º 54/2015, p. 3, 24 mar. 2015.

_____. **Resolução nº 60, de 19 de setembro de 2008**. Institui o Código de Ética da Magistratura Nacional. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/127>. Acesso em: 09 dez. 2021.

_____. Resolução nº 61, de 07 de outubro de 2008. Disciplina o procedimento de cadastramento de conta única para efeito de constrição de valores em dinheiro por intermédio do Convênio BACENJUD e dá outras providências. **Diário da Justiça**: Brasília, DF, n.º 68/2008, p. 2-4, 15 out. 2008.

_____. **Resolução nº 106, de 06 de abril de 2010**. Dispõe sobre os critérios objetivos para aferição do merecimento para promoção de magistrados e acesso aos Tribunais de 2º grau. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/168>. Acesso em: 11 jun. 2022.

_____. **Resolução nº 331, de 20 de agosto de 2020**. Institui a Base Nacional de Dados do Poder Judiciário – DataJud como fonte primária de dados do Sistema de Estatística do Poder Judiciário – SIESPJ para os tribunais indicados nos incisos II a VII do art. 92 da Constituição Federal. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3428>. Acesso em: 12 set. 2022.

_____. Resolução nº 332, de 21 de agosto de 2020. Dispõe sobre a ética, a transparência e a governança na produção e no uso de Inteligência Artificial no Poder Judiciário e dá outras providências. **Diário da Justiça**: Brasília, DF, n.º 274, p. 4-8, 25 ago. 2020.

_____. **Resolução nº 446, de 14 de março de 2022**. Institui a plataforma Codex como ferramenta oficial de extração de dados estruturados e não estruturados dos processos judiciais eletrônicos em tramitação no Poder Judiciário Nacional e dá outras providências. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/4417>. Acesso em: 12 set. 2022.

_____. **TJSC firma acordo e implantará sistema eletrônico do CNJ**. [2020]. Disponível em: <https://www.cnj.jus.br/tjsc-firma-acordo-e-implantara-sistema-eletronico-do-cnj>. Acesso: 01 mai. 2020.

CONSELHO NACIONAL DE JUSTIÇA; CONSELHO SUPERIOR DA JUSTIÇA DO TRABALHO. **Acordo de Cooperação Técnica nº 01/2010**. Acordo de Cooperação Técnica que entre si celebram o Tribunal Superior do Trabalho, o Conselho Superior da Justiça do Trabalho e os Tribunais Regionais do Trabalho. Brasília: CNJ e CSJT, 29 mar. 2010. Disponível em: <https://www.csjt.jus.br/documents/955023/9626845/Desenvolvimento+do+sistema+d>

e+processo+judicial+eletr%C3%B4nico.pdf/d54be507-cabc-73c5-ab7c-2dd8d07a8c97?t=1635361011200. Acesso em: 02 fev. 2022.

CONSANI, Cristina Foroni. A Crítica de Jeremy Waldron ao Constitucionalismo Contemporâneo. **Revista da Faculdade de Direito – UFPR**, Curitiba, v. 59, n. 2, p. 143-173, 2014.

COSTA, Ramon Silva; OLIVEIRA, Samuel Rodrigues. Pode a Máquina Julgar? Considerações Sobre o Uso de Inteligência Artificial no Processo de Decisão Judicial. **Revista de Argumentação e Hermenêutica Jurídica**, v. 4, n. 2, p. 21-39, jul/dez. 2018.

CROOTOFF, Rebecca. 'Cyborg Justice' and the Risk of Technological-Legal Lock-In. **Columbia Law Review Forum**, New York, v. 119, p. 233-251, 20 nov. 2019.

CRUZ, Fabricio Bittencourt da; BRITTO, Melina Carla de Souza; MOREIRA, Guilherme Martelli; BRITTO JUNIOR, Alceu de Souza. Robôs Substituem Juízes? O Estado da Arte da Inteligência Artificial no Judiciário Brasileiro. **Revista Antinomias**, v. 3, n. 1, p. 8-41, jan./jun., 2022.

DALL·E. **The Artist's Garden at Giverny, Monet**. 17 jun. 2022. Imagem PNG. 1024 pixels x 1024 pixels.

D'AMATO, Antony. Legal Uncertainty. **California Law Review**, Berkeley, v. 71, n. 1, p. 1-55, January, 1983. DOI: 10.2307/3480139.

DANZIGER, Shai; LEVAV, Jonathan; Avnaim-Pesso, Liora. Extraneous factors in judicial decisions. **PNAS**, v. 108, n. 17, April 26, 2011. DOI: 10.1073/pnas.1018033108.

DEEKS, Ashley. The Judicial Demand for Explainable Artificial Intelligence. . **Columbia Law Review**, v. 119, n. 7, p. 1829-1850, 2020.

DEEPMIND. **AlphaCode**. [2022]. Disponível em: <https://alphacode.deepmind.com>. Acesso em: 04 fev. 2022.

DENNETT, Daniel C. **From Bacteria to Bach and Back**. London: Penguin Books, 2017. E-book Kindle.

DEWEY, Daniel; RUSSEL, Stuart; TEGMARK, Max. Research Priorities for Robust and Beneficial Artificial Intelligence. **AI Magazine**, v. 36, n. 4, 2015.

DIDÁTICA Tech. **O que são Redes Neurais e Deep Learning**. [2022]. Disponível em: <https://didatica.tech/introducao-a-redes-neurais-e-deep-learning>. Acesso em: 31 jan. 2023.

DI BELLO, Marcello; VERHEIJ, Bart. Evidence & decision making in the law: theoretical, computational and empirical approaches. **Artificial Intelligence and Law**, Dordrecht, v. 28, 1-5, 2020. DOI: 10.1007/s10506-019-09253-0.

DICK, Philip K. **The Minority Report**: And Other Classic Stories. New York: Citadel Press, 2002.

DINAMARCO, Cândido Rangel; LOPES, Bruno Vasconcelos Carrilho. **Teoria Geral do Novo Processo Civil**. 2ª ed. São Paulo: Malheiros, 2017.

DIOUF, Rabiyaou; SARR, Edouard Ngor; SALL, Ousmane; BIRREGAH, Babiga; BOUSSO, Mamadou; MBAYE, Sény Ndiaye. Web Scraping: State-of-the-Art and Areas of Application. **2019 IEEE International Conference on Big Data (Big Data)**, Los Angeles, p. 6040-6042, 2019. DOI: 10.1109/BigData47090.2019.9005594.

DIXON, Rosalind. Argumento Central a Favor da Forma Fraca do Controle de Constitucionalidade. **Revista Direitos Fundamentais & Democracia**, Curitiba, v. 24, n. 2, p. 5-55, mai./ago. 2019. DOI: 10.25192/issn.1982-0496.rdfd.v24i21647.

DONEDA, Danilo. A proteção dos dados pessoais como um direito fundamental. **Espaço Jurídico Journal of Law**, v. 12, n. 2, p. 91-108, 13 dez. 2011.

DOWNEY, Allen B. **Pense em Python**: pense como um cientista da computação. Tradução: Sheila Gomes. São Paulo: Novatec, 2016.

DOWNS, Anthony. Teoría económica de la acción política en una democracia. In: RUBIO, Albert Batlle (coord.). **Diez textos básicos de Ciencia Política**. Madrid: Ariel, 2001.

DUNAGAN, Jake. Neuro-Futures: The Brain, Politics, and Power. **Journal of Futures Studies**, Taipei, v. 9, n. 2, p. 1-18, 2004.

DWORKIN, Ronald. **A Raposa e o Porco-espinho**: Justiça e Valor. Tradução: Marcelo Brandão Cipolla. São Paulo: WMF Martins Fontes, 2014.

_____. **Levando os Direitos a Sério**. 3ª ed. Tradução: Nelson Boeira. São Paulo: Martins Fontes, 2010.

_____. **O Império do Direito**. Tradução de Jefferson Luiz Camargo. São Paulo: Martins Fontes, 1999.

_____. **O Império do Direito**. Tradução: Jeferson Luiz Camargo. 3ª ed. São Paulo: Martins Fontes, 2014.

EDGAR, Craig. Struggling Against Entropy: Monetary Patent Infringement Damages after eBay. **Drake Law Review**, Des Moines, v. 66, p. 45-67, 2018.

EKMAN, Paul. **Emotions Revealed**: Understanding Faces and Feelings. London: Orion. E-book Kindle.

ELIOT, Lance B. **AI and Legal Reasoning Essentials**. Stanford: LBE Press, 2020.

ELSTER, Jon. **Sour Grapes**: Studies in the Subversion of Rationality. Cambridge: Cambridge University Press, 1983. E-book Kindle.

_____. Ulysses and the sirens: A theory of imperfect rationality. **Social Science Information**, v. 16, n. 5, 469-526, 1977. DOI: 10.1177/053901847701600501.

_____. **Ulysses Unbound: Studies in Rationality, Precommitment, and Constraints.** Cambridge: Cambridge University Press, 2000. E-book Kindle.

ESCOVEDO, Tatiana; KOSHIYAMA, Adriano. **Introdução a Data Science:** Algoritmos de Machine Learning e métodos de análise. São Paulo: Casa do Código, 2020, posição 3030. E-book Kindle.

ESTADOS UNIDOS DA AMÉRICA. **AI Principles: Recommendations on the Ethical Use of Artificial Intelligence by the Department of Defense.** [2019]. Disponível em: https://media.defense.gov/2019/Oct/31/2002204458/-1/-1/0/DIB_AI_PRINCIPLES_PRIMARY_DOCUMENT.PDF. Acesso em: 08 ago. 2020.

_____. **Blueprint for an AI Bill Of Rights.** [2022]. Disponível em: <https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf>. Acesso em: 27 dez. 2022.

EVANS, Eric. **Domain-Driven Design:** Atacando as Complexidades no Coração do Software. 3ª ed. Tradução: Tibério Júlio Couto Novais. Rio de Janeiro: Alta Books, 2016.

FENOLL, Jordi Nieva. **Inteligencia Artificial y Proceso Judicial.** Madri: Marcial Pons, 2018.

FERNANDES, Ivo André Domingues. **A Deep Learning Approach to Named Entity Recognition in Portuguese Texts.** 2010. Dissertação (Mestrado Integrado em Engenharia Informática e Computação). Faculdade de Engenharia da Universidade do Porto (FEUP), Porto, 2018.

FERNANDES, W. P. D., SILVA, L. J. S., FRAJHOF, I. Z. *et al.* Appellate Court Modifications Extraction for Portuguese. **Artificial Intelligence and Law**, 2019. DOI: 10.1007/s10506-019-09256-x.

FERRARA, Massimiliano; GAGLIOTI, Roberto Angelo. Legal Values and Legal Entropy: A Suggested Mathematical Model. **International Journal Of Mathematical Models And Methods In Applied Sciences**, v. 6, n. 3, p. 490-498, 2012.

FERRUCCI, David et al. Watson: Beyond Jeopardy! In: **Artificial Intelligence**, 199-200, 2013, p. 93-105.

FLORIDI, Luciano. Information Ethics, Its Nature and Scope. **SIGCAS Computer and Society**, v. 36, n. 3, Set. 2006.

FORST, Rainer. **Justificação e Crítica:** perspectivas de uma teoria crítica da política. Tradução: Denilson Luis Werle. São Paulo: Editora Unesp Digital, 2019. E-book Kindle.

FRIEDRICH, Roland; LUZZATTO, Mauro; ASH, Elliott. Entropy in Legal Language. **CEUR Workshop Proceedings**, Germany, v. 2645, p. 25-30, ago. 2020. DOI: 10.3929/ethz-b-000427800.

FROMER, Jeanne C. An Information Theory of Copyright Law. **Emory Law Journal**, v. 64, n. 1, p. 71-128, 2014.

FUNDAÇÃO GETÚLIO VARGAS. **Supremo em Números: Plataforma de Referências.** [2021]. Disponível em: http://precedentes.cloud.fgv.br/search?termo_busca=razoabilidade. Acesso em: 31 jul. 2021.

FUTURE OF LIFE INSTITUTE. **Asilomar AI Principles.** [2017]. Disponível em: <https://futureoflife.org/ai-principles>. Acesso em: 05 ago. 2020.

G20. **G20 AI Principles.** [2019]. Disponível em: <https://www.g20-insights.org/wp-content/uploads/2019/07/G20-Japan-AI-Principles.pdf>. Acesso em: 25 jul. 2020.

G7. **G7's Charlevoix Common Vision for the Future of AI.** [2018]. Disponível em: <https://www.mofa.go.jp/mofaj/files/000373837.pdf>. Acesso em: 25 jul. 2020.

GADAMER, Hans-Georg. **Hermenêutica em Retrospectiva.** Tradução: Marco Antônio Casanova. Petrópolis: Vozes, 2009.

_____. **Verdade e Método:** Traços fundamentais de uma hermenêutica filosófica. 3ª ed. Tradução: Flávio Paulo Meurer. Petrópolis: Editora Vozes, 1999.

_____. **Verdade e Método:** traços fundamentais de uma hermenêutica filosófica. 4.ª ed. Petrópolis: Vozes, 2002.

GALLES, David. **Comparison Sorting Algorithms.** [2020]. Disponível em: <https://www.cs.usfca.edu/~galles/visualization/ComparisonSort.html>. Acesso em: 06 jan. 2022.

GANMOTE, Shubham V.; VISHWAS G. R. S.; KUMAR, Anupama. Analysis of Sorting Algorithms Using Time Complexity. **International Journal of Engineering Research & Technology (IJERT)**, Gandhinagar, v. 5, n. 21, 2017.

GARGARELLA, Roberto. **As Teorias da Justiça Depois de Rawls:** Um breve manual de filosofia política. Tradução: Alonso Reis Freire. São Paulo: WMF Martins Fontes, 2008.

_____. **La Justicia Frente al Gobierno.** Quito: Corte Constitucional para el Período de Transición, 2011.

GELL-MAN, Murray. **El Quark y el Jaguar:** Aventuras en lo simple y lo complejo. Traducción: Ambrosio García Leal y Romualdo Pastor. Barcelona: Tusquets Editores, 1995.

GILL, Richard David; FENTON, Norman; LAGNADO, David; NEIL, Martin. Modelling Competing Legal Arguments Using Bayesian Model Comparison and Averaging. **Artificial Intelligence and Law**, v. 27, p. 403-430, 2019. DOI: 10.1007/s10506-019-09250-3.

GILPIN, Leilani H. et. al. Explaining Explanations: An Overview of Interpretability of Machine Learning. **2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)**, Turin, Italy, p. 80-89, 2018. DOI: 10.1109/DSAA.2018.00018.

GITHUB. **GitHub Colpilot: Your AI pair programmer**. [2021]. Acesso em: 04 fev. 2022.

GOOGLE Cloud. **Produtos de IA e machine learning**. [2022]. Disponível em: <https://cloud.google.com/products/ai>. Acesso em 15 nov. 2022.

GRAFFIN, Greg. I Love My Computer. In: BAD Religion. **The New America**. Nova York: Atlantic Records, 2000. Disponível em: <https://youtu.be/zIM0vahvauU>. Acesso em: 01 jun. 2022.

GRECO, Luís. **Poder de Julgar Sem Responsabilidade de Julgador**: impossibilidade jurídica do juiz-robô. São Paulo: Marcial Pons, 2020.

GRECO, Rodrigo Azevedo. **Direito e Entropia**. Fortaleza: do autor, 2015. E-book Kindle.

GREENFIELD, Adam. **Everyware**: The dawning age of ubiquitous computing. Berkeley: Pearson Education, 2006. E-book Kindle.

GRUS, Joel. **Data Science do Zero**. Tradução: Welington Nascimento. Rio de Janeiro: Alta Books, 2016.

GUIDOTTI, Riccardo; MONREALE, Anna; RUGGIERI, Salvatore; TURINI, Franco; GIANNOTTI, Fosca; PEDRESCHI, Dino. A Survey of Methods for Explaining Black Box Models. **ACM Computing Surveys**, v. 51, n. 5, Article 93, January 2019, 42 pages. DOI: 10.1145/3236009.

GUSTIN, Miracy Barbosa de Sousa; DIAS, Maria Tereza Fonseca; NICÁCIO, Camila Silva. **(Re)pensando a Pesquisa Jurídica**: Teoria e Prática. São Paulo: Almedina, 2020. E-book Kindle.

HABERMAS, Jürgen. **A Inclusão do Outro**: Estudos de teoria política. São Paulo: Loyola, 2002.

_____. **Direito e Democracia**: entre facticidade e validade. Vol. I. Tradução de Flávio Beno Siebeneichler. Brasília: Tempo Brasileiro, 1997.

_____. La Lucha por el Reconocimiento en el Estado Democrático de Derecho. In: TAYLOR, Charles; et al. **El Multiculturalismo y "La Política del Reconocimiento"**. Tradução: Mónica Utrilla de Neira, Liliana Andrade Llanas e Gerard Vilar Roca. Ciudad de México: Fondo de Cultura Económica, 2009. E-book Kindle.

_____. **Teoria do Agir Comunicativo, 1**: Racionalidade da Ação e Racionalidade Social. Tradução: Paulo Astor Soethe. São Paulo: WMF Martins Fontes, 2012.

HARARI, Yuval Noah. **Homo Deus**: Uma breve história do amanhã. Tradução: Paulo Geiger. São Paulo: Companhia das Letras, 2016.

_____. **Sapiens**: Uma breve história da humanidade. Porto Alegre: L&PM Editores, 2015. E-book Kindle.

HART, Herbert L. A. **O Conceito de Direito**. Tradução: Antônio de Oliveira Sette-Câmara. São Paulo: WMF Martins Fontes, 2009.

HARTMANN PEIXOTO, Fabiano. **Inteligência Artificial e Direito**: Convergência Ética e Estratégica. Vol. 5. Curitiba: Alteridade, 2020.

_____. Projeto Victor: Relato do Desenvolvimento da Inteligência Artificial na Repercussão Geral do Supremo Tribunal Federal. **Revista Brasileira de Inteligência Artificial e Direito**, v. 1, n. 1, jan-abr. 2020.

HARTMANN PEIXOTO, Fabiano; SILVA, Roberta Zumblick Martins da. **Inteligência Artificial e Direito**. Vol. 1. Curitiba: Alteridade, 2019.

HAWKING, Stephen. **Uma Breve História do Tempo**. Tradução: Cássio de Arantes Leite. Rio de Janeiro: Intrínseca, 2015.

HEGEL, Georg Wilhelm Friedrich. **Princípios da Filosofia do Direito**. Tradução: Orlando Vitorino. São Paulo: Martins Fontes, 1997.

HEIDEGGER, Martin. **Ser e Tempo**. Tradução: Márcia Sá Cavalcante Schuback. 4ª ed. Petrópolis: Vozes, 2009.

HOMMERDING, Adalberto Narciso; HARTMANN, Gabriel Henrique. Direito, Hermenêutica e Inteligência Artificial: Construindo Pontes entre Decisão Judicial, Compreensão Existencial e Mecanismos de Machine Learning. **Quaestio Iuris**, v. 14, n. 4, Rio de Janeiro, p. 1001-1038, 2021. DOI: 10.12957/rqi.2021.58084.

HOWARD, Philip N. **Pax Technica**: How the Internet of Things May Set Us Free or Lock Us Up. New Haven: Yale University Press, 2015. E-book Kindle.

IBÁÑEZ, Perfecto Andrés. ¿Neutralidad o Pluralismo en La Aplicación Del Derecho? Interpretación Judicial e Insuficiencia Del Formalismo. **Doxa**, Madri, n. 15-16, p. 861-882, 1994.

IBM. **Build with Watson: The AI for business**. [2020]. Disponível em: <https://www.ibm.com/watson/developer>. Acesso em: 09 ago. 2020.

_____. **Soluções de IA**. [2022]. Disponível em: <https://www.ibm.com/br-pt/artificial-intelligence>. Acesso em 15 nov. 2022.

INGLATERRA. England and Wales Court of Appeal. Associated Provincial Picture Houses Ltd. vs. Wednesbury Corporation, 1948, 1 KB 223, p. 228-230.

ISMAIL, Salim; VAN GEES, Yuri; MALONE, Michael S. **Organizações exponenciais**. Rio de Janeiro: Alta Books, 2018. E-book Kindle.

JOTA. **Judiciário brasileiro tem ao menos 72 projetos de inteligência artificial nos tribunais**: Ferramentas são usadas para auxiliar agrupamento de demandas repetitivas a até para sugerir minutas. [2020]. Disponível em: <https://www.jota.info/coberturas-especiais/inoa-e-acao/judiciario-brasileiro-tem-ao-menos-72-projetos-de-inteligencia-artificial-nos-tribunais-09072020>. Acesso em: 20 ago. 2020.

KAFRI, Oded; KAFRI, Hava. **Entropy: God's Dice Game**. Scotts Valley: CreateSpace, 2013. E-book Kindle.

KAHNEMAN, Daniel. **Rápido e Devagar**: Duas formas de pensar. Tradução: Cássio de Arantes Leite. Rio de Janeiro: Objetiva, 2012. E-book Kindle.

KAHNEMAN, Daniel; SIBONY, Olivier; SUNSTEIN, Cass R. **Noise: A Flaw in Human Judgment**. New York: Little, Brown Spark, 2021. E-book Kindle.

KAHNEMAN, Daniel; TVERSKY, Amos. **Judgement Under Uncertainty: Heuristics and Biases**. Cambridge: Cambridge University Press, 1982. E-book Kindle.

KANT, Immanuel. **Crítica da Razão Prática**. Tradução: Monique Hulshof. Petrópolis: Editora Vozes; Bragança Paulista: Editora Universitária São Francisco, 2017.

_____. **Crítica da Razão Pura**. Edição do Kindle, 2020.

KATZ, Daniel Martin; BOMMARITO, MJ. Measuring the complexity of the law: the United States Code. **Artificial Intelligence and Law**, Dordrecht, v. 22, p. 337-374, 16 sep. 2014. DOI: 10.1007/s10506-014-9160-8.

KELSEN, Hans. **Teoria Geral do Direito e do Estado**. Tradução: Luís Carlos Borges. 4ª ed. São Paulo: Martins Fontes, 2005.

_____. **Teoria Pura do Direito**. Tradução de João Baptista Machado. São Paulo: Martins Fontes, 1998.

_____. **Teoria Pura do Direito**. Tradução: João Baptista Machado. 8ª ed. São Paulo: WMF Martins Fontes, 2009.

KING JR, Joseph H. Deus ex Machina and the Unfulfilled Promise of New York Times v. Sullivan: Applying the Times for All Seasons. **Kentucky Law Journal**, Limestone Lexington, v. 95, n. 3, p. 650-674, 2007.

KOSTER, Raphael; BALAGUER, Jan; TACCHETTI, Andrea et al. Human-centred mechanism design with Democratic AI. **Nature Human Behaviour**, v. 6, p. 1398-1407, 2022. DOI: 10.1038/s41562-022-01383-x.

KURZWEIL, Ray. **A Singularidade Está Próxima**: Quando os Humanos Transcendem a Biologia. Tradução: tradução Ana Goldberger. São Paulo: Iluminuras, 2018.

_____. **Como Criar uma Mente**: os segredos do pensamento humano. São Paulo: Aleph, 2014. E-book Kindle.

KYMLICKA, Will. **Contemporary Political Philosophy: An Introduction**. 2ª ed. Oxford: Oxford University Press, 2002.

_____. **Filosofia Política Contemporânea**. Tradução: Luís Carlos Borges. São Paulo: Martins Fontes, 2006.

LAGE, Fernanda de Carvalho. **Manual de Inteligência Artificial no Direito Brasileiro**. Salvador: Editora Juspodivm, 2021.

LEAL, Rogério Gesta; SWAROVSKY, Aline. As Garantias Constitucionais do Processo no Brasil. **Revista de Direitos Fundamentais & Democracia**, Curitiba, v. 9, n. 9, p. 242-277, 2011.

LEGRAND, Pierre. **Como Ler o Direito Estrangeiro**. Tradução: Daniel Wunder Hachem. São Paulo: Contracorrente, 2018. E-book Kindle.

LEVITSKY, Steven; ZIBLAT, Daniel. **Como as Democracias Morrem**. Rio de Janeiro: Zahar, 2018.

LI, N.; LIU, S.; LIU, Y.; ZHAO, S.; LIU, M. Neural Speech Synthesis with Transformer Network. **Proceedings of the AAI Conference on Artificial Intelligence**, v. 33, n. 1, p. 6706-6713, 2019. DOI: 10.1609/aaai.v33i01.33016706.

LI, Yujia et al. **Competition-Level Code Generation with AlphaCode**. Disponível em: https://storage.googleapis.com/deepmind-media/AlphaCode/competition_level_code_generation_with_alphacode.pdf. Acesso em: 04 fev. 2022.

LICHT, K. de Fine; LICHT, J. de Fine. Artificial intelligence, transparency, and public decision-making. **AI & Society**, Mar. 2020. DOI: 10.1007/s00146-w.

LIN, David Chuan-En. **8 Simple Techniques to Prevent Overfitting**. [2020]. Disponível em: <https://towardsdatascience.com/8-simple-techniques-to-prevent-overfitting-4d443da2ef7d>. Acesso em: 31 jan. 2023.

LINNA JR., Daniel W. What We Know and Need to Know About Legal Startups. **South Carolina Law Review**, V. 67, p. 389-417, 2016.

LOPEZ, Martin Andreoni; SANZ, Igor Jochem; LOBATO, Antonio G. P. Aprendizado de Máquina em Plataformas de Processamento Distribuído de Fluxo: Análise e Detecção de Ameaças em Tempo Real. **Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos (SBRC) - Minicursos**, maio 2018. Disponível em: <http://143.54.25.88/index.php/sbrccminicursos/article/view/1768>. Acesso em 28 nov. 2022.

LORENZETTO, Bruno Meneses. **O Silêncio das Sereias: Tempo, Direito e Violência na Modernidade**. Dissertação (Mestrado em Direito). Universidade Federal do Paraná (UFPR), Curitiba, 2010.

_____. **Os Caminhos do Constitucionalismo Para a Democracia**. Tese (Doutorado em Direito). Universidade Federal do Paraná (UFPR), Curitiba, 2014.

LORENZETTO, Bruno Meneses; CLÈVE, Clèmerson Merlin. Interpretação Constitucional: entre dinâmica e integridade. **Seqüência**, Florianópolis, n. 72, p. 67-92, abr. 2016. DOI: <http://dx.doi.org/10.5007/2177-7055.2016v37n72p67>.

_____; _____. Constituição Federal, Controle Jurisdicional e Níveis de Escrutínio. **Direitos Fundamentais & Justiça**, Curitiba, v. 32, p. 97-123, jul./set. 2015.

_____; _____. Teorias Interpretativas, Capacidades Institucionais e Crítica. **Revista Direitos Fundamentais & Democracia**, Curitiba, v. 19, n. 19, jan./jun. 2016.

MACCORMICK, Neil. **Argumentação Jurídica e Teoria do Direito**. Tradução: Waldéa Barcellos. 2ª ed. São Paulo: WMF Martins Fontes, 2009.

_____. **Retórica e o Estado de Direito**: Uma teoria da argumentação jurídica. Tradução: Conrado Hübner Mendes e Marcos Paulo Veríssimo. Rio de Janeiro: Elsevier, 2008.

_____. **Retórica y Estado de Derecho**: Una teoría del razonamiento jurídico. Lima: Palestra Editores, 2017. E-book Kindle.

MAGRANI, Eduardo. **A Internet das Coisas**. Rio de Janeiro: FGV Editora, 2018.

_____. **Entre Dados e Robôs**: Ética e Privacidade da Era da Hiperconectividade. 2ª ed. Porto Alegre: Arquipélago Editorial, 2019.

MALISKA, Marcos Augusto. **Fundamentos da Constituição**: abertura, cooperação, integração. Curitiba: Juruá, 2013.

MARANHÃO, Juliano Souza de Albuquerque; FLORÊNCIO, Juliana Abrusio; ALMADA, Marco. Inteligência artificial aplicada ao direito e direito da inteligência artificial. **SUPREMA**: Revista de Estudos Constitucionais, Brasília, v. 1, n. 1, p. 154-180, jan/jun. 2021.

MARINONI, Luiz Guilherme. **Precedentes Obrigatórios**. 2ª ed. São Paulo: Revista dos Tribunais 2011.

MARQUES, Cláudia Lima. Superação das Antinomias pelo Diálogo das Fontes: O Modelo Brasileiro de Coexistência entre o Código de Defesa do Consumidor e o Código Civil de 2002. **Revista da Escola Superior da Magistratura de Sergipe**, Aracajú, n. 07, 2004.

MARRAFON, Marco Aurélio. **O Caráter Complexo da Decisão em Matéria Constitucional**: Discursos sobre a Verdade, Radicalização Hermenêutica e Fundação Ética na Práxis Jurisdicional. Rio de Janeiro: Lumen Juris, 2010.

MARTIN, Robert C. **Código Limpo**: Habilidades Práticas do Agile Software. Tradução: Leandro Chu. Rio de Janeiro: Alta Books, 2011.

MATHEWS, Jud; SWEET, Alec Stone. All Things in Proportion? American Rights Doctrine and the Problem of Balancing. **Emory Law Journal**, v. 60, p. 798-875, 2011.

MAYOR, Adrienne. **Gods and Robots**: Myths, Machines and Ancient Dreams of Technology. Princeton: Princeton University Press, 2018. E-book Kindle.

MEDAUAR, Odete. **Direito Administrativo Moderno**. 19ª ed. São Paulo: Revista dos Tribunais, 2015.

MEDVEDEVA, Masha; VOLS, Michel; WIELING, Martijn. Using machine learning to predict decisions of the European Court of Human Rights. **Artificial Intelligence and Law**, Dordrecht, 2019. DOI: 10.1007/s10506-019-09255-y.

MENDES, Gilmar Ferreira. Parcerias Institucionais e o Conselho Nacional de Justiça: instrumento de efetivação do acesso à Justiça. **Revista Jurídica da Presidência**, Brasília, v. 20, n. 122, p. 516-540, Out. 2018/Jan. 2019. DOI: 10.20499/2236-3645.RJP2018v20e122-1819.

MENDES, Laura Schertel Ferreira. In: DONEDA, Danilo; SARLET, Ingo Wolfgang; MENDES, Laura Schertel Ferreira. **Estudos Sobre Proteção de Dados Pessoais: Dados num Mundo em Transformação**. São Paulo: Expressa, 2022. E-book Kindle.

MERKL, Adolf Julius. **Escritos de Teoria do Direito**. Organização e Tradução: Matheus Pelegriño da Silva. São Leopoldo: Editora UNISINOS, 2018.

MIAILLE, Michel. **Introdução Crítica ao Direito**. 3ª ed. Tradução: Ana Prata. Lisboa: Editorial Estampa, 2005.

MICHELMAN, Frank. Foreword: Traces of Self-Government. **Harvard Law Review**, Cambridge, v. 100, n. 1, p. 4-77, nov. 1986.

MICROSOFT. **Azure AI**. [2022]. Disponível em: <https://azure.microsoft.com/pt-pt/solutions/ai>. Acesso em 15 nov. 2022.

MINAS GERAIS. Tribunal de Justiça do Estado de Minas Gerais. **TJMG utiliza inteligência artificial em julgamento virtual**. [2020]. Disponível em: <https://www.tjmg.jus.br/portal-tjmg/noticias/tjmg-utiliza-inteligencia-artificial-em-julgamento-virtual.htm>. Acesso em: 15 set. 2020.

MIT Technology Review. **Por que a GPT-3 é o melhor e o pior da IA atualmente**. [2021]. Disponível em: <https://mittechreview.com.br/por-que-a-gpt-3-e-o-melhor-e-o-pior-da-ia-atualmente>. Acesso em: 08 abr. 2021.

MONET, Claude. **Le Jardin de l'artiste à Giverny**. 1900. Óleo sobre tela. 84,6 cm x 92,6 cm.

MOORE, Gordon E. Cramming more components onto integrated circuits. In: **Electronics**, vol. 38, n. 8, April 19, 1965.

MORAES, José Diniz de; TABAK, Benjamin Miranda. As Heurísticas e Vieses da Decisão Judicial: análise econômico-comportamental do direito. **Revista Direito GV**, São Paulo, v. 14, n. 2, p. 618-653, set. 2018. DOI: 10.1590/2317-6172201824.

MORAN, Jay P. Postmodernism's Misguided Place in Legal Scholarship: Chaos Theory, Deconstruction, and Some Insights from Thomas Pynchon's Fiction. **Southern California Interdisciplinary Law Journal**, Los Angeles, p. 155-199, 1997.

MOTA, Caio et al. Named Entity Recognition in Portuguese Legal Documents Using Deep Learning. In: **Anais do XVIII Encontro Nacional de Inteligência Artificial e Computacional**, Porto Alegre, p. 130-140, 2021. DOI: 10.5753/eniac.2021.18247.

MOZEIC, Vinícius Almada. Os sistemas jurídicos inteligentes e o caminho perigoso até a teoria da argumentação de Robert Alexy. **Revista Brasileira de Direito**, Passo Fundo, v. 13, n. 3, 2017.

MÜLLER, Friedrich. **Metodologia do Direito Constitucional**. 4.^a ed. rev. e atual. e ampl. Tradução: Peter Naumann. São Paulo: Revista dos Tribunais, 2010.

_____. **Teoria Estruturante do Direito**. 2.^a ed. rev., atual. e amp. Tradução: Peter Naumann e Eurides Avance de Souza. São Paulo: Revista dos Tribunais, 2009.

NEWTON, Isaac. **Letter from Sir Isaac Newton to Robert Hooke**, 1675. Disponível em: <https://digitallibrary.hsp.org/index.php/Detail/objects/9792>. Acesso em: 06 abr. 2021.

NGUYEN, T., NGUYEN, L., TOJO, S. *et al.* Recurrent neural network-based models for recognizing requisite and effectuation parts in legal texts. **Artificial Intelligence and Law**, v. 26, p. 169–199, 2018. DOI: 10.1007/s10506-018-9225-1.

NIETZSCHE, Friedrich. **Humano, Demasiado Humano**. Tradução: Paulo César de Souza. São Paulo: Companhia das Letras, 2000. E-book Kindle.

NISHIMOTO, Shinji; VU, An T.; NASELARIS, Thomas; BENJAMINI, Yuval; YU, Bin; GALLANT, Jack L. Reconstructing Visual Experiences from Brain Activity Evoked by Natural Movies. **Current Biology**, v. 21, p. 1641-1646, Berkeley, 2011. DOI: 10.1016/j.cub.2011.08.031.

NUMFOCUS Inc. **Numpy documentation**. [2022]. Disponível em: <https://numpy.org/doc>. Acesso em 15 nov. 2022.

_____. **Pandas documentation**. [2022]. Disponível em: <https://pandas.pydata.org/docs>. Acesso em 15 nov. 2022.

NUNES, Marcelo Guedes. **Jurimetria: como a estatística pode reinventar o Direito**. 2.^a ed. rev. e atual. São Paulo: Revista dos Tribunais, 2019.

OCDE. **Recommendation 0449, of 21 May 2019: Recommendation of the Council on Artificial Intelligence**. Disponível em: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>. Acesso em: 04 ago. 2020.

O'NEIL, Cathy. **Weapons of Math Destruction: how Big Data increases inequality and threatens democracy**. New York: Crown, 2016. E-book Kindle.

OPENAI. **ChatGPT**. [2022]. Disponível em: <https://chat.openai.com/chat>. Acesso em 20 dez. 2022.

_____. **DALL·E**. [2022]. Disponível em: <https://labs.openai.com>. Acesso em 12 jun. 2022.

_____. E-mail de OpenAI para Sérgio Rodrigo de Pádua com o título: **“Welcome to the private beta for our API!”**, 2021. Acesso em 21 dez. 2022.

_____. **GPT-4 can solve difficult problems with greater accuracy, thanks to its broader general knowledge and problem solving abilities.** [2023]. Disponível em: <https://openai.com/product/gpt-4>. Acesso em: 20 mar. 2023.

_____. **OpenAI API.** [2021]. Disponível em: <https://openai.com/blog/openai-api/>. Acesso em 08 abr. 2021.

PÁDUA, Sérgio Rodrigo de. O Juiz Ciborgue: Inteligência Artificial e Decisão Judicial. In: ROVER, Aires José; et al. (coord.). **Direito, Governança e Novas Tecnologias III.** Florianópolis: CONPEDI, 2020, p. 315-335.

PARANÁ. Tribunal de Justiça do Estado do Paraná. **Pedido de Providências 0103379-12.2022.8.16.6000.** Reunião TJPR-UnB realizada em 18/08/2022 (Vídeo nº 8063985 e Ata nº 8067385). Curitiba, PR: Tribunal de Justiça do Estado do Paraná.

_____. Tribunal de Justiça do Estado do Paraná. Resolução nº 03, de 14 de maio de 2009. Dispõe sobre o processo eletrônico no âmbito do Poder Judiciário do Estado do Paraná. **Diário da Justiça Eletrônico:** Curitiba, PR, n.º 138, 14 mai. 2009.

_____. Tribunal de Justiça do Estado do Paraná. **TJPR vai lançar novos sistemas de inteligência artificial.** [2019]. Disponível em: <https://www.tjpr.jus.br/noticias>. Acesso em: 01 set. 2019.

PARISI, Francesco. Entropy in Property. **The American Journal of Comparative Law**, v. 50, n. 3, p. 595-632, Summer 2002. DOI: 10.2307/841063.

PASQUALE, Frank. Data-Informed Duties in AI Development. **Columbia Law Review**, v. 119, n. 7, p. 1917-1940, 2019.

_____. **The Black Box Society:** The Secret Algorithms That Control Money and Information. Cambridge: Harvard University Press, 2015. E-book Kindle.

_____. Toward a Fourth Law of Robotics: Preserving Attribution, Responsibility, and Explainability in an Algorithmic Society. **Ohio State Law Journal**, Vol. 78, 2017.

PEIRIS, G. L. Wednesbury Unreasonableness: The Expanding Canvas. **Cambridge Law Journal**, v. 46, n. 1, p. 53-82, March 1987.

PEREIRA, João Sérgio dos Santos Soares. **A Padronização Decisória na Era da Inteligência Artificial:** uma possível leitura hermenêutica e da autonomia do direito. Belo Horizonte: Casa do Direito, 2021.

PEREIRA, Sebastião Tavares. O *Machine Learning* e o Máximo Apoio Ao Juiz. **Revista Democracia Digital e Governo Eletrônico**, Florianópolis, v. 1, n. 18, p. 2-35, 2019.

PINTO, Cristiano Paixão Araujo. **Modernidade, Tempo e Direito.** Belo Horizonte: Del Rey, 2002.

PIRES, Teresinha Inês Teles. Teorias da Argumentação Jurídica e a Prática Discursiva do Supremo Tribunal Federal: Metodologias Analíticas e Estudo de Casos

(AC 4.070/DF e ADI 5526/DF). **Revista Direitos Fundamentais & Democracia**, Curitiba, v. 24 n. 2, p. 102-127, 2019.

POPPER, Karl. **A Lógica da Pesquisa Científica**. Tradução: Leonidas Hegenberg e Octanny Silveira da Mota. 2ª ed. São Paulo: Cultrix, 2013.

POURABDOLLAH, Amir; BRAILSFORD, Tim. On the Turing Completeness of the Semantic Web. **Theory of Computing Systems**, v. 56, 2014, p. 1-18. DOI: 10.1007/s00224-014-9536-x.

POZEN, David E. Transparency's Ideological Drift. **The Yale Law Journal**, v. 128, n.1, p. 100-165, 2018.

PRAKKEN, Henry. AI & Law, Logic and Argument Schemes. **Argumentation**, v. 19, p. 303-320, 2005. DOI: 10.1007/s10503-005-4418-7.

PRATA, Raquel. O Juiz Cyborg - A Informática e o Direito. **Revista Julgar**, Lisboa, n. 8, 2009.

PRIGOGINE, Ilya. **O Fim das Certezas**: Tempo, Caos e as Leis da Natureza. Tradução: Roberto Leal Ferreira. 2ª ed. São Paulo: Editora UNESP, 2011.

PROVOST, Foster; FAWCETT, Tom. **Data Science para Negócios**. Rio de Janeiro: Alta Books, 2016.

PUGLIESI, William Soares. **Princípios da Jurisprudência**. Belo Horizonte: Arraes Editores, 2017.

PYCARET. **PyCaret Oficial**. [2022]. Disponível em: <https://pycaret.gitbook.io/docs>. Acesso em 15 nov. 2022.

PYTHON Software Foundation. **Classes**. [2022]. Disponível em: <https://docs.python.org/pt-br/3/tutorial/classes.html>. Acesso em 15 nov. 2022.

_____. **Datetime**. [2022]. Disponível em: <https://docs.python.org/pt-br/3/library/datetime.html>. Acesso em 15 nov. 2022.

QUACH, Katyanna. **A developer built an AI chatbot using GPT-3 that helped a man speak again to his late fiancée. OpenAI shut it down**. [2021]. Disponível em: https://www.theregister.com/2021/09/08/project_december_openai_gpt_3/. Acesso em 03 jan. 2023.

RAMOS, Janine Vilas Boas Gonçalves. **Inteligência Artificial no Poder Judiciário Brasileiro**: Projetos de IA nos tribunais e o sistema de apoio ao processo decisório judicial. São Paulo: Dialética, 2022. E-book Kindle.

RAWLS, John. **A Theory of Justice**. Cambridge: Harvard University Press, 1971.

_____. **Justiça e Democracia**. Tradução: Irene A. Paternot. São Paulo: Martins Fontes, 2000.

_____. **O Liberalismo Político**. 2ª ed. Tradução: Dinah de Abreu Azevedo São Paulo: Ática, 2000.

RAZ, Joseph. **O Conceito de Sistema Jurídico**: Uma Introdução à Teoria dos Sistemas Jurídicos. Tradução: Maria Cecília Almeida. São Paulo: WMF Martins Fontes, 2012.

_____. Pode Haver Uma Teoria do Direito? In: RAZ, Joseph. **Uma Discussão Sobre Teoria do Direito**. Tradução: Sheila Stolz. São Paulo: Marcial Pons, 2013.

REED, Scott *et al.* A Generalist Agent. **Transactions on Machine Learning Research**. [2022]. Disponível em: <https://openreview.net/pdf?id=1ikK0kHvj>. Acesso em: 16 nov. 2022.

ROBL FILHO, Ilton Norberto. **Accountability e Independência Judiciais**: O Desenho Institucional do Judiciário e do Conselho Nacional de Justiça no Estado Democrático de Direito Brasileiro. 2012. Tese (Doutorado em Direito). Universidade Federal do Paraná (UFPR), Curitiba, 2012.

RONDÔNIA. Tribunal de Justiça do Estado de Rondônia. **Inteligência Artificial desenvolvida pelo TJRO pode revolucionar o Judiciário**. [2018]. Disponível em: <https://www.tjro.jus.br/noticias/item/9472-inteligencia-artificial-desenvolvida-pelo-tjro-pode-revolucionar-o-judiciario>. Acesso em: 01 set. 2019.

ROSE, Janus. **Facebook Pulls Its New ‘AI For Science’ Because It’s Broken and Terrible**. [2022]. Disponível em: <https://www.vice.com/en/article/3adyw9/facebook-pulls-its-new-ai-for-science-because-its-broken-and-terrible>. Acesso em 29 nov. 2022.

ROSS Intelligence. **Announcement**. [2021]. Disponível em: blog.rossintelligence.com/post/announcement. Acesso em: 04 nov. 2021.

_____. **Ross Intelligence**. [2019]. Disponível em: <https://rossintelligence.com>. Acesso em: 31 ago. 2019.

ROVELLI, Carlo. **A Ordem do Tempo**. Tradução: Silvana Cobucci. Rio de Janeiro: Objetiva, 2017.

_____. **A Realidade Não é o Que Parece**: a estrutura elementar das coisas. Tradução: Silvana Cobucci Leite. Rio de Janeiro: Objetiva, 2014. E-book Kindle.

ROVER, Aires José. **Informática no Direito**: Inteligência Artificial. Curitiba: Juruá, 2001.

_____. *Machine Learning* no Poder Judiciário: uma biblioteca temática, ano 2020. In: HARTMANN PEIXOTO, Fabiano (Org.). **Inteligência Artificial**: estudos de inteligência artificial. Vol. 4. Curitiba: Alteridade, 2021.

_____. Sistemas Especialistas Legais: Pensando o Sistema Jurídico. **Sequência, estudos jurídicos e políticos**, Florianópolis, v. 29, p. 65-79, dez. 1994.

RUNCIMAN, David. **Como a Democracia Chega ao Fim**. Tradução: Sergio Flaksman. São Paulo: Todavia, 2018.

RUSSELL, Stuart; NORVIG, Peter. **Inteligência Artificial**. Tradução de Regina Célia Simille. Rio de Janeiro: Elsevier, 2013.

SALOMÃO, Luis Felipe (coord.) et al. **Inteligência Artificial: Tecnologia Aplicada à Gestão dos Conflitos no Âmbito do Poder Judiciário Brasileiro**. 2ª ed. FGV Conhecimento: Centro de Inovação, Administração e Pesquisa do Judiciário, 2022. Disponível em: https://ciapj.fgv.br/sites/ciapj.fgv.br/files/relatorio_ia_2fase.pdf. Acesso em: 01 jun. 2022.

SÁNCHEZ-CUENCA, Ignacio. Power, Rules, and Compliance. In: MARAVALL, José María; PRZEWORSKI, Adam (eds.). **Democracy and the Rule of Law**. Cambridge: Cambridge University Press, 2003.

SANTA CATARINA. Tribunal de Justiça de Santa Catarina. **Processo Eletrônico SAJ: Histórico**. [2020]. Disponível em: <https://www.tjsc.jus.br/web/processo-eletronico-saj/historico>. Acesso em 01 mai. 2020.

SÃO PAULO. Tribunal de Justiça do Estado do Paraná. Resolução nº 551, de 31 de agosto de 2011. Regulamenta o processo eletrônico no âmbito do Tribunal de Justiça do Estado de São Paulo e dá outras providências. **Diário da Justiça Eletrônico**: São Paulo, SP, Ano IV, edição n.º 1033, 08 set. 2011.

SARLET, Gabrielle Bezerra Sales; RUARO, Regina Linden. A Proteção de Dados Sensíveis no Sistema Normativo Brasileiro sob o Enfoque da Lei Geral de Proteção de Dados (LGPD) – L.13.709/2018. **Revista Direitos Fundamentais & Democracia**, Curitiba, v. 26, n. 2, p. 81-106, mai./ago. 2021. DOI: 10.25192/issn.1982-0496.rdfd.v26i22172.

SARTOR, Giovanni. Doing justice to rights and values: teleological reasoning and proportionality. **Artificial Intelligence and Law**, Dordrecht, v. 18, p. 175-215, 2010. DOI: 10.1007/s10506-010-9095-7.

_____. Legal Concepts as Inferential Nodes and Ontological Categories. **Artificial Intelligence and Law Review**, Dordrecht, v. 17, p. 217-251, 2009. DOI: 10.1007/s10506-009-9079-7.

SCHAUER, Frederick. **Playing by the Rules: A Philosophical Examination of Rule-Based Decision-Making I Law and Life**. Oxford: Oxford University Press, 1991. E-book Kindle.

SCHILDT, Herbert. **Java para iniciantes: crie, compile e execute programas Java rapidamente**. Tradução: Aldir José Coelho Corrêa da Silva. 6ª ed. Porto Alegre: Bookman, 2015.

SCHWAB, Klaus. **A Quarta Revolução Industrial**. Tradução: Daniel Moreira Miranda. São Paulo: Edipro, 2018. E-book Kindle.

SCIKIT-LEARN. **Scikit-learn: Machine Learning in Python**. [2022]. Disponível em: <https://scikit-learn.org>. Acesso em 15 nov. 2022.

SEN, Amartya. **A Ideia de Justiça**. Tradução: Denise Bottmann e Ricardo Doninelli Mendes. São Paulo: Companhia das Letras, 2009. E-book Kindle.

SHANNON, Claude E. A Mathematical Theory of Communication. **The Bell System Technical Journal**, v. 27, n. 3, p. 379-423, Jul. 1948. DOI: 10.1002/j.1538-7305.1948.tb01338.x.

SHANNON, Claude E.; WEAVER, Warren. **The Mathematical Theory of Communication**. Chicago: University of Illinois Press, 1963. E-book Kindle.

SHEARER, Colin. The CRISP-DM Model: The New Blueprint for Data Mining. **Journal of Data Warehousing**, v. 5, n. 4, 2000.

SICHELMAN, Ted. Quantifying Legal Entropy. **Frontiers in Physics**, Lausanne, v. 9, article 665054, 21 jun. 2021. DOI: 10.3389/fphy.2021.665054.

SILVA, Diego Mattozo Bernardes da. **Métodos de categorização de variáveis preditoras em modelos de regressão para variáveis binárias**. 2017. Dissertação (Mestrado em Estatística). Universidade Federal de São Carlos (UFScar), São Carlos, 2017.

SOLUM, Lawrence B. Originalism *Versus* Living Constitutionalism: The Conceptual Structure of the Great Debate. **Northwestern University Law Review**, Chicago, v. 113, n. 6, p. 1243-1296, 2019.

SOURDIN, Tania. Judge v Robot? Artificial Intelligence and Judicial Decision-Making. **UNSW Law Journal**, Sydney, v. 41, n. 4. p. 1114-1133, 2018.

SPOREA, Daniel-Cosmin. Judecătorul Cyborg: O Automatizare a (In)Justiției Penale? **Revista Universul Juridic**, București, n.1, p. 71-79, 2022.

STEIBEL, Fabro; VICENTE, Victor; JESUS, Diego Santos Vieira de. Possibilidades e Potenciais da Utilização da Inteligência Artificial. In: FRAZÃO, Ana (coord.) et. al. **Inteligência Artificial e Direito: Ética, Regulação e Responsabilidade**. São Paulo: Thomson Reuters Brasil, 2019.

STEPHAN, Paul B. One Voice in Foreign Relations and Federal Common Law. **Virginia Journal of International Law**, v. 60, p. 1-79, 2019.

STOPFORD, Ben. **Designing Event-Driven Systems: Concepts and Patterns for Streaming Services with Apache Kafka**. Sebastopol: O'Reilly, 2018.

STRECK, Lenio; BERNSTES, Luísa Giuliani; GOMES, Jefferson de Carvalho. Inteligência Artificial: mesmos problemas, mas na versão Hi-Tech. **Constituição, Economia e Desenvolvimento: Revista da Academia Brasileira de Direito Constitucional**, Curitiba, 2021, vol. 13, n. 25, p. 333-342, ago./dez., 2021.

SUNDFELD, Carlos Ari. **Direito Administrativo para Céticos**. São Paulo: Malheiros, 2012.

SUNSTEIN, Cass R. Algorithms, Correcting Biases. **Social Research: Algorithms**, New York, v. 86, n. 2, 2019.

_____. **Choosing Not to Choose**. Oxford: Oxford University Press, 2015. E-book Kindle.

SUSSKIND, Richard. **Online Courts and the Future of Justice**. Oxford: Oxford University Press, 2019. E-book Kindle.

_____. **The End of Lawyers?** Rethinking the Nature of Legal Services. Oxford: Oxford University Press, 2010. E-book Kindle.

_____. **Tomorrow's Lawyers: An Introduction to Your Future**. 2ª ed. Oxford: Oxford University Press, 2017. E-book Kindle.

TALEB, Nassim N. **A Lógica do Cisne Negro**: O Impacto do Altamente Improvável. Tradução Marcelo Schild. 1ª ed. Rio de Janeiro: BestBusiness, 2019.

TAULLI, Tom. **Introdução à Inteligência Artificial**: Uma abordagem não técnica. São Paulo: Novatec, 2020.

TAYLOR, Charles. **A Secular Age**. Cambridge: Harvard University Press, 2007. E-book Kindle.

_____. **As Fontes do Self**: A construção da identidade moderna. Tradução: Adail Ubirajara Sobral e Dinah de Abreu Azevedo. 4ª ed. São Paulo: Edições Loyola, 2013.

_____. La Política del Reconocimiento. In: TAYLOR, Charles; et al. **El Multiculturalismo y "La Política del Reconocimiento"**. Tradução: Mónica Utrilla de Neira, Liliana Andrade Llanas e Gerard Vilar Roca. Ciudad de México: Fondo de Cultura Económica, 2009. E-book Kindle.

TAYLOR, Ross *et al.* **Galactica**: A Large Language Model for Science. [2022]. Disponível em: <https://galactica.org/static/paper.pdf>. Acesso em: 16 nov. 2022.

THALER, Richard H.; SUNSTEIN, Cass R. **Nudge**: como tomar melhores decisões sobre saúde, dinheiro e felicidade. Tradução: Ângelo Lessa. Rio de Janeiro: Objetiva, 2019. E-book Kindle.

THE LINUX Foundation. **PyTorch**. [2022]. Disponível em: <https://pytorch.org/docs/stable/index.html>. Acesso em 15 nov. 2022.

THE MATRIX Reloaded. Direção: de Lana Wachowski e Lilly Wachowski. Produção: Warner Bros. EUA: 2003.

TURING, Alan Mathison. Computing Machinery and Intelligence. **Mind**, n. 49, p. 433-460, 1950.

_____. On computable numbers, with an application to the *Entscheidungs* problem. **Journal of Math**, v. 58, 1936, p. 345-363.

TUSHNET, Mark. **I Dissent**: Great Opposing Opinions in Landmark Supreme Court Cases. Boston: Beacon Press, 2008. E-book Kindle.

UNESCO. **Recommendation on the Ethics of Artificial Intelligence**. 24 Nov 2021. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000380455>. Acesso em: 20 mar. 2022.

UNIÃO EUROPEIA. **Carta de Direitos Fundamentais da União Europeia**. Disponível em: https://www.europarl.europa.eu/charter/pdf/text_pt.pdf. Acesso em: 03 jan. 2023.

_____. **Diretiva (UE) 2016/680 do Parlamento Europeu e do Conselho, de 27 de abril de 2016**. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:32016L0680&from=HR>. Acesso em: 05 jan. 2023.

_____. **Ethics Guidelines for Trustworthy AI**. [2019]. 08 April 2019. Disponível em: https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419. Acesso em: 25 jul. 2020.

_____. **Proposta de Diretiva do Parlamento Europeu e do Conselho Relativa à Adaptação das Regras de Responsabilidade Civil Extracontratual à Inteligência Artificial**. [2022]. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:52022PC0496&from=EN>. Acesso em: 27 dez. 2022.

_____. **Regulamento (UE) 2016/679 do Parlamento Europeu e do Conselho, de 27 de abril de 2016**. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:32016R0679&from=PT>. Acesso em: 05 jan. 2023.

VALENTINI, Rômulo Soares. **Julgamento por Computadores? As Novas Possibilidades da Juscibernética no Século XXI e Suas Implicações Para o Futuro do Direito e do Trabalho dos Juristas**. 2017. Tese (Doutorado em Direito). Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, 2017.

VAN WYK, Jo-Ansie. South Africa's Nuclear Diplomacy Since The Termination of Its Nuclear Weapons Programme. **Scientia Militaria - South African Journal of Military Studies**, v. 42, p. 80-101. DOI: 10.5787/42-1-1082.

VASCONCELLOS, Fernando Andreoni. **Hermenêutica Jurídica e Derrotabilidade**. Curitiba: Juruá, 2010.

VEGA, Italo S. Inteligência Artificial e Tomada de Decisão – A necessidade de agentes externos. In: FRAZÃO, Ana (coord.) et. al. **Inteligência Artificial e Direito: Ética, Regulação e Responsabilidade**. São Paulo: Thomson Reuters Brasil, 2019.

VON NEUMANN, John. **The Computer and the Brain**. New Haven: Yale University Press, 1958.

WAGNER, Ben. Accountability by design in technology research. **Computer Law & Security Review**, v. 37, 105398, ISSN 0267-3649, Oxford, 2020. DOI: 10.1016/j.clsr.2020.105398.

WALDRON, Jeremy. **Law and Disagreement**. New York: Oxford, 2004.

_____. The Core of the Case Against Judicial Review. **Yale Law Journal**, v. 115, n. 6, 2006.

WALKER, Joshua. **On Legal AI: Um Rápido Tratado sobre a Inteligência Artificial no Direito**. Tradução: José Ignacio Coelho Mendes Neto. São Paulo: Revista dos Tribunais, 2021.

WALTON, Douglas. **Argumentation Methods for Artificial Intelligence in Law**. Berlin: Springer, 2005. E-book Kindle.

WANG, Daniel Wei Liang. From Wednesbury Unreasonableness to Accountability for Reasonableness. **Cambridge Law Journal**, v. 76, n. 3, p. 642-670, 2017.

WARAT, Luis Alberto. **O Direito e Sua Linguagem**. Porto Alegre: Sérgio Fabris, 1984.

_____. Saber Crítico e Senso Comum Teórico dos Juristas. **Seqüência: Estudos Jurídicos e Políticos**, Florianópolis, v. 3, n. 5, p. 48-57, 1982. DOI: 10.5007/%25x.

WEBER, Max. **Economia e Sociedade: Fundamentos da sociologia compreensiva**. Vol. 2. Tradução: Regis Barbosa e Karen Elsabe Barbosa. Brasília: Editora Universidade de Brasília, 1999.

WITTGENSTEIN, Ludwig. **Investigações Filosóficas**. Tradução: José Carlos Bruni. São Paulo: Nova Cultural, 1999.

_____. **Observações Filosóficas**. Tradução: Adail Sobral e Maria Stela Gonçalves. São Paulo: Edições Loyola, 2005.

_____. **Tractatus Logico-Philosophicus**. Tradução: Luiz Henrique Lopes dos Santos. São Paulo: Editora Universidade de São Paulo, 2020.

YAPO, Adrienne; WEISS, Joseph. Ethical Implications of Bias in Machine Learning. **Proceedings of 51st Hawaii International Conference on System Sciences**, Waikoloa Village, 2018.

ZUBOFF, Shoshana. **A Era do Capitalismo de Vigilância: A luta por um futuro humano na nova fronteira de poder**. Tradução: George Schlesinger. Intrínseca: Rio de Janeiro, 2019. E-book Kindle.