XV ENCONTRO INTERNACIONAL DO CONPEDI ALICANTE - ESPANHA

GOVERNO DIGITAL, DIREITO E NOVAS TECNOLOGIAS II

FELIPE CHIARELLO DE SOUZA PINTO

FABIANO HARTMANN PEIXOTO

Copyright © 2026 Conselho Nacional de Pesquisa e Pós-Graduação em Direito

Todos os direitos reservados e protegidos. Nenhuma parte destes anais poderá ser reproduzida ou transmitida sejam quais forem os meios empregados sem prévia autorização dos editores.

Diretoria – CONPEDI
Presidente - Profa. Dra. Samyra Haydêe Dal Farra Naspolini - FMU – São Paulo
Diretor Executivo - Prof. Dr. Orides Mezzaroba - UFSC – Santa Catarina
Vice-presidente Norte - Prof. Dr. Jean Carlos Dias - Cesupa – Pará
Vice-presidente Centro-Oeste - Prof. Dr. José Querino Tavares Neto - UFG – Goiás
Vice-presidente Sul - Prof. Dr. Leonel Severo Rocha - Unisinos – Rio Grande do Sul
Vice-presidente Sudeste - Profa. Dra. Rosângela Lunardelli Cavallazzi - UFRJ/PUCRio – Rio de Janeiro
Vice-presidente Nordeste - Prof. Dr. Raymundo Juliano Feitosa - UNICAP - Pernambuco

Representante Discente: Prof. Dr. Abner da Silva Jaques - UPM/UNIGRAN - Mato Grosso do Sul

Conselho Fiscal:
Prof. Dr. José Filomeno de Moraes Filho - UFMA - Maranhão
Prof. Dr. Caio Augusto Souza Lara - SKEMA/ESDHC/UFMG - Minas Gerais
Prof. Dr. Valter Moura do Carmo - UFERSA - Rio Grande do Norte
Prof. Dr. Fernando Passos – UNIARA - São Paulo
Prof. Dr. Edinilson Donisete Machado - UNIVEM/UENP - São Paulo
Secretarias
Relações Institucionais:
Prof. Dra. Claudia Maria Barbosa - PUCPR - Paraná
Prof. Dr. Heron José de Santana Gordilho - UFBA - Bahia
Profa. Dra. Daniela Marques de Moraes - UNB - Distrito Federal
Comunicação:
Prof. Dr. Robison Tramontina - UNOESC - Santa Catarina
Prof. Dr. Liton Lanes Pilau Sobrinho - UPF/Univali - Rio Grande do Sul
Prof. Dr. Lucas Gonçalves da Silva - UFS - Sergipe
Relações Internacionais para o Continente Americano:
Prof. Dr. Jerônimo Siqueira Tybusch - UFSM - Rio Grande do Sul
Prof. Dr. Paulo Roberto Barbosa Ramos - UFMA - Maranhão
Prof. Dr. Felipe Chiarello de Souza Pinto - UPM – São Paulo
Relações Internacionais para os demais Continentes:
Profa. Dra. Gina Vidal Marcilio Pompeu - UNIFOR - Ceará
Profa. Dra. Sandra Regina Martini - UNIRITTER / UFRGS - Rio Grande do Sul
Profa. Dra. Maria Claudia da Silva Antunes de Souza - UNIVALI – Santa Catarina
Educação Jurídica
Profa. Dra. Viviane Coêlho de Séllos Knoerr - Unicuritiba - PR
Prof. Dr. Rubens Beçak - USP - SP
Profa. Dra. Livia Gaigher Bosio Campello – UFMS – MS
Eventos:
Prof. Dr. Yuri Nathan da Costa Lannes - FDF - São Paulo
Profa. Dra. Norma Sueli Padilha - UFSC - Santa Catarina
Prof. Dr. Juraci Mourão Lopes Filho - UNICHRISTUS - Ceará

Comissão Especial
Prof. Dr. João Marcelo de Lima Assafim - UFRJ - RJ
Profa. Dra. Maria Creusa De Araújo Borges - UFPB - PB
Prof. Dr. Antônio Carlos Diniz Murta - Fumec - MG
Prof. Dr. Rogério Borba - UNIFACVEST - SC

G326

Governo digital, direito e novas tecnologias I[Recurso eletrônico on-line] organização CONPEDI

Coordenadores: Yuri Nathan da Costa Lannes; José Renato Gaziero Cella; Cinthia Obladen de Almendra Freitas– Florianópolis: CONPEDI, 2026.

Inclui bibliografia
ISBN: 978-65-5274-659-7
Modo de acesso: www.conpedi.org.br em publicações
Tema: Meio ambiente e sustentabilidade: perspectivas globais e locais na agenda 2030

2. Governo digital. 3. Direito e novas tecnologias. XV Encontro Internacional do CONPEDI Alicante - Espanha (1: 2026: Florianópolis, Brasil).

CDU: 34



Conselho Nacional de Pesquisa
e Pós-Graduação em Direito Florianópolis
Santa Catarina – Brasil
www.conpedi.org.br

XV ENCONTRO INTERNACIONAL DO CONPEDI
ALICANTE - ESPANHA

GOVERNO DIGITAL, DIREITO E NOVAS TECNOLOGIAS II

Apresentacao

GT - Governo Digital, Direito e Novas Tecnologias II

O Grupo de Trabalho - GT Governo Digital, Direito e Novas Tecnologias II, ocorreu em 29 de maio de 2026, no XV Encontro Internacional do Conpedi - Alicante - Espanha, cujo tema principal foi Meio Ambiente e Sustentabilidade: perspectivas globais e locais na Agenda 2030.

Mantendo a sua tradição, o GT reuniu diversos trabalhos na sua área temática, a seguir resumidamente apresentados pelos coordenadores: Felipe Chiarello de Souza Pinto, Fabiano Hartmann Peixoto e Angela Issa Haonat.

Apresentação:

A SOCIEDADE DE ALGORITMOS E NÃO-COISAS NA FORMAÇÃO DE UM CIBERESPAÇO DE PODER

O artigo foi apresentado por Cinthia Obladen de Almendra Freitas. Com base na noção de infosfera de Luciano Floridi, compreendida como o continuum que integra humanos e tecnologias, e na ideia de “não-coisas” de Byung-Chul Han, o artigo aponta para tensões a partir de um ciberespaço que necessita ser reinterpretado.

INFRAESTRUTURA DIGITAL, TERRITÓRIO E REGULAÇÃO: PROPOSTA DE “VALUTAÇÃO” DE IMPACTO TERRITORIAL DIGITAL (VITD) APLICADA A DATA CENTERS

O artigo foi apresentado por Nicolò Basigli, com a presença do coautor Valter Moura do Carmo. O artigo analisa a materialidade territorial da infraestrutura digital, a localização de data centers e a distribuição territorial dos custos ambientais associados à economia de dados, relacionando essas questões à literatura sobre justiça ambiental e à regulação pública da infraestrutura tecnológica. A pesquisa adota abordagem qualitativa, com base em revisão bibliográfica interdisciplinar e análise normativa, examinando instrumentos regulatórios existentes no contexto europeu e brasileiro.

A REGULAÇÃO DA INTELIGÊNCIA ARTIFICIAL NO FEDERALISMO BRASILEIRO: REPARTIÇÃO DE COMPETÊNCIAS, GOVERNANÇA MULTINÍVEL E ATUAÇÃO DOS ENTES SUBNACIONAIS

O artigo foi apresentado por Leonardo David Quintiliano e analisa a regulação da inteligência

artificial no contexto do federalismo brasileiro, com foco na repartição constitucional de competências entre União, Estados, Distrito Federal e Municípios. Parte-se da constatação de que a expansão das aplicações de IA impõe a construção de um marco regulatório capaz de conciliar inovação tecnológica, segurança jurídica e proteção de direitos fundamentais. Inicialmente, examinaram-se modelos internacionais de governança da IA, com destaque para os Estados Unidos, a China e a União Europeia, especialmente o Artificial Intelligence Act, e concluiu-se que a regulação da IA no Brasil tende a assumir feição multinível, exigindo coordenação federativa e cooperação institucional.

A GEOPOLÍTICA DA INTELIGÊNCIA ARTIFICIAL E A SUSTENTABILIDADE DOS DATACENTERS: DESAFIOS REGULATÓRIOS PARA O SUL GLOBAL NA ERA DO COLONIALISMO DIGITAL.

O artigo apresentado por Debora Bonat, com a presença dos coautores, mostra que a aparente imaterialidade da Inteligência Artificial oculta uma infraestrutura física massiva de datacenters com consumo alarmante de água e energia elétrica. Essa dinâmica gera um "colonialismo digital", no qual Big Techs do Norte Global exploram recursos naturais do Sul Global enquanto retêm os lucros e o controle tecnológico. Diante disso, legislações atuais como o AI Act europeu e o PL 2.338/2023 no Brasil mostram-se insuficientes por focarem quase exclusivamente em privacidade e vieses algorítmicos. Para reverter esse cenário, defende-se a urgência de um Direito Digital ecologicamente orientado que integre métricas socioambientais rígidas às avaliações de impacto e promova a transição para a Green AI.

PROCEDIMENTALIZAÇÃO DO RACIOCÍNIO JURÍDICO EM SISTEMAS DE INTELIGÊNCIA ARTIFICIAL: ARQUITETURAS HÍBRIDAS E EXPLICABILIDADE

O artigo apresentado por Izabela Frota Melo, com a presença dos coautores, analisa os desafios do uso de grandes modelos de linguagem (LLMs) no direito, destacando sua capacidade de gerar respostas plausíveis, mas sem fundamentação jurídica transparente. Defende que arquiteturas híbridas, que combinam IA generativa e modelos formais de argumentação, podem tornar as decisões mais explicáveis, analisáveis e contestáveis. Sustenta que a formalização jurídica não desaparece, mas assume novas funções de orientação, reconstrução e avaliação das justificações produzidas pela IA.

A CAPTURA DA JUSTIÇA PELOS INDICADORES: LIMITES E RISCOS DA GESTÃO POR DESEMPENHO NO PODER JUDICIÁRIO BRASILEIRO

O artigo foi apresentado por Lia Loana Curial Oliva , José Renato Gaziero Cella , Orides Mezzaroba e analisou criticamente o modelo de governança pública adotado pelo Conselho Nacional de Justiça (CNJ), orientado por indicadores de produtividade e seus impactos sobre a concepção e a efetividade da prestação jurisdicional no Brasil. Propôs reorientar a governança judicial para incorporar métricas qualitativas que permitam avaliar a qualidade da prestação jurisdicional e articulou fundamentos normativos e epistemológicos da governança com uma crítica filosófica inspirada em Kafka, revelando os paradoxos éticos da quantificação da justiça no contexto brasileiro.

ENTRE O ALGORITMO E A TOGA: USO RESPONSÁVEL DE IA GENERATIVA NO

JUDICIÁRIO E OS DESAFIOS DO DEVIDO PROCESSO TECNOLÓGICO

O artigo foi apresentado por José Renato Gaziero Cella, com coautor, e examinou os riscos do uso irresponsável de inteligência artificial generativa no exercício da função jurisdicional, tendo como fio condutor o caso do Tribunal de Justiça de Minas Gerais (TJMG) em que um prompt de ferramenta generativa foi inadvertidamente deixado no corpo de acórdão que absolveu acusado de estupro de vulnerável contra criança de 12 anos. O artigo apresentou a conclusão de que a IA pode e deve auxiliar o Judiciário, mas jamais substituir o núcleo valorativo do ato de julgar: a responsabilidade pela fundamentação permanece integralmente humana.

OPACIDADE ALGORÍTMICA E LEGITIMIDADE DAS DECISÕES ADMINISTRATIVAS AUTOMATIZADAS: UMA ANÁLISE A PARTIR DA TEORIA DISCURSIVA DO DIREITO DE ROBERT ALEXY

O artigo foi apresentado por Djônata Winter e demonstrou que os requisitos de transparência, explicabilidade e motivação das decisões administrativas automatizadas não constituem meras exigências técnicas ou éticas, mas pressupostos conceituais e normativos da própria validade jurídica desses atos, à luz da Teoria Discursiva do Direito de Robert Alexy. Concluiu que transparência, explicabilidade e motivação constituem o equivalente funcional da argumentação racional no contexto tecnológico, condicionando a legitimidade do uso de sistemas de IA pela Administração Pública à possibilidade de reconstrução discursiva de seus procedimentos e resultados.

ÁGUA PARA OS ALGORITMOS: CONFLITOS SOCIOAMBIENTAIS ENTRE DATA CENTERS E COMUNIDADES TRADICIONAIS NO BRASIL

O artigo apresentado por Nicolò Basigli abordou que a expansão da Inteligência Artificial tem sido frequentemente apresentada como um fenômeno essencialmente tecnológico e imaterial, associado à crescente digitalização da sociedade, mas que o funcionamento desses sistemas depende de uma vasta infraestrutura material responsável pelo processamento e armazenamento de dados em escala global. Assim, o artigo analisou a relação entre Inteligência Artificial, infraestrutura digital e recursos naturais, com especial atenção ao papel dos data centers na economia contemporânea de dados.

IA E ELEIÇÕES 2026: LABORATÓRIO DE IA E INTEGRIDADE ELEITORAL (DR.IA-AIEL) NA MICROSEGMENTAÇÃO GENERATIVA DO DISCURSO POLÍTICO DIGITAL.

O artigo apresentado por Fabiano Hartmann relata a possibilidade de irmos além da interferência da IA no processo eleitoral. Fora do campo das fake news, o artigo analisou a possibilidade de combinação de sistema de IA para microsegmentação política e análise de sentimentos, que, combinados com o conceito de "general influence" permitem a geração de linguagem argumentativa de acordo com cada microsegmento eleitoral. O artigo ainda encaminha para a necessidade de desenvolvimento de conhecimento e regulação sobre a temática, tendo em vista o impacto potencial no processo eleitoral.

CONSTITUCIONALISMO DIGITAL E REGULAÇÃO DAS PLATAFORMAS NA ORDEM ECONÔMICA.

O artigo apresentado por Jesualdo Eduardo de Almeida Júnior analisou a reconfiguração da esfera pública promovida pelas plataformas digitais, investigando em que medida tais atores passaram a desempenhar funções materialmente públicas sem estarem submetidos a um regime jurídico compatível com sua relevância institucional.

TECNOCRACIA, TECNOPÓLIO E PODER INFORMACIONAL: REFLEXÕES JURÍDICAS SOBRE A CENTRALIZAÇÃO TECNOLÓGICA NA SOCIEDADE DIGITAL

O artigo apresentado por Jesualdo Eduardo de Almeida Junior aborda as profundas transformações nas estruturas de poder da sociedade contemporânea pela expansão das tecnologias digitais. Assim, o artigo analisou os impactos jurídicos da centralização tecnológica na sociedade digital, investigando como a concentração de dados, infraestruturas e sistemas algorítmicos produz novas formas de poder informacional e concluiu que a centralização tecnológica impõe ao direito o desenvolvimento de novas estratégias regulatórias voltadas à proteção da autonomia informacional, à transparência algorítmica e à preservação da legitimidade democrática no ambiente digital.

A TRANSNACIONALIDADE NA REGULAÇÃO DA INTELIGÊNCIA ARTIFICIAL

O artigo apresentado por Djônata Winter apresenta a constatação de um cenário em que o desenvolvimento e a aplicação da Inteligência Artificial transcendem as fronteiras nacionais e desafiam as estruturas regulatórias tradicionais baseadas em territorialidade, Este artigo analisa a transnacionalidade na regulação dessa tecnologia, mapeando seus atores, instrumentos, desafios e tendências.

Por fim, a professora Angela Issa Haonat, que estava auxiliando também na coordenação do GT, apresentou seu artigo na temática do acesso à justiça e identidade de gênero. Desafios da prestação jurisdicional e do atendimento institucional à população trans no sistema brasileiro de justiça. Relatou que seu trabalho foi desenvolvido em coautoria com Maria Clara Rezende Duarte Queiroz. O trabalho relata as pesquisas junto à comissão de gênero e analisa o acesso à justiça da população trans partindo de uma premissa crítica, segundo a qual, embora o Estado Democrático de Direito preveja universalidade, grupos vulnerabilizados ainda enfrentam barreiras institucionais, simbólicas e estruturais do sistema de justiça brasileiro. A investigação teve como objeto o aprimoramento das práticas institucionais para ampliar esse acesso e a efetividade dos direitos dessa população. Concluiu que não basta o Provimento 73 do CNJ, que dispõe sobre a verbação da alteração do prenome do gênero nos assentamentos de pessoas transgênero no registro civil ou decisões do STF. É preciso que as rotinas administrativas e os fluxos de atendimento reflitam o respeito à identidade auto-percebida.

Os diálogos oportunizados no GT certamente contribuirão para o avanço das pesquisas, registrando, os coordenadores, o alto nível dos trabalhos apresentados e o caminho profícuo do GT na trajetória do CONPEDI.

Os coordenadores,

Procedimentalização do Raciocínio Jurídico em Sistemas de Inteligência Artificial: Arquiteturas Híbridas e Explicabilidade

Proceduralization of Legal Reasoning in Artificial Intelligence: Hybrid Architectures and Explainability

Izabela Frota Melo
Debora Bonat

Resumo

O surgimento de grandes modelos de linguagem transformou a inteligência artificial aplicada ao direito ao permitir a produção de interpretações e justificações jurídicas plausíveis sem formalização normativa explícita. Esse movimento desloca o problema central da área: da modelagem formal do raciocínio jurídico para a explicitação e o controle das justificações geradas por sistemas neurais.

O artigo examina esse deslocamento a partir de arquiteturas híbridas que combinam modelos neurais de geração textual com estruturas formais de representação e análise argumentativa. Sustenta-se que a formalização do raciocínio jurídico, nesse contexto, passa a desempenhar três funções: estruturar o raciocínio decisório, orientar a geração de justificações e reconstruir e avaliar a coerência justificatória das respostas produzidas.

Argumenta-se que essas estruturas permitem submeter tais respostas a critérios jurídicos de justificação, controle e contestação, aproximando a explicabilidade de uma condição de legitimidade das decisões assistidas por inteligência artificial.

Conclui-se que esses sistemas não substituem a interpretação jurídica institucional, mas procedimentalizam a produção e o controle das justificações jurídicas, contribuindo para o dever de fundamentação e para a efetividade do direito à explicação em contextos de decisão automatizada.

Palavras-chave: Inteligência Artificial Jurídica, Grandes Modelos de Linguagem, Arquiteturas Híbridas, Explicabilidade, Raciocínio Jurídico, Direito à Explicação

Abstract/Resumen/Résumé

The emergence of large language models has transformed artificial intelligence in law by enabling the production of plausible legal interpretations and justifications without explicit normative formalization. This shift redefines the field’s central problem: from the formal modeling of legal reasoning to the explicitation and control of justifications generated by

neural systems.

This paper examines this transformation through hybrid architectures that combine neural text-generation models with formal structures for representation and argumentative analysis. It argues that, in this context, the formalization of legal reasoning performs three functions: structuring decision-making, guiding the generation of justifications, and reconstructing and evaluating justificatory coherence.

It is further argued that these structures allow such outputs to be subjected to legal criteria of justification, control, and contestation, bringing explainability closer to a condition of legitimacy for AI-assisted decisions.

The article concludes that these systems do not replace institutional legal interpretation but proceduralize the production and control of legal justifications, contributing to the duty to give reasons and to the effectiveness of the right to explanation in automated decision-making contexts.

Keywords/Palabras-claves/Mots-clés: Legal Artificial Intelligence, Large Language Models, Hybrid Architectures, Explainability, Legal Reasoning, Right to Explanation

1. Introdução

O uso de sistemas de inteligência artificial na prática jurídica tem se expandido rapidamente nos últimos anos. Entre as tecnologias mais recentes, destacam-se os grandes modelos de linguagem (Large Language Models – LLMs), capazes de produzir textos complexos em linguagem natural e de responder questões jurídicas com elevado grau de fluidez discursiva. Esses sistemas podem sintetizar precedentes, articular argumentos normativos e apresentar justificações extensas para soluções jurídicas.

Esse desenvolvimento tecnológico, entretanto, introduz um problema para o campo da inteligência artificial aplicada ao direito. Embora tais sistemas revelem capacidade de produzir respostas juridicamente plausíveis, a estrutura de fundamentação subjacente permanece, em grande medida, opaca. Essa opacidade é agravada pela própria natureza dos dados de treino que alimentam tais modelos, extraídos de grandes conjuntos de dados não curados da internet, os quais incorporam e reproduzem visões dominantes e vieses estruturais, incluindo associações estereotipadas relacionadas a gênero, raça, etnia e deficiência (Bender et al., 2021).

Diferentemente de sistemas anteriores, nos quais o raciocínio decisório era explicitamente formalizado por meio de regras ou estruturas formais de argumentação, os sistemas de aprendizado de máquina, especialmente aqueles baseados em aprendizado profundo (deep learning), operam por meio de processos estatísticos de geração textual, sem formalização normativa explícita.

Embora possam oferecer apoio inicial à redação de decisões judiciais, com aparência de coerência e de adequada fundamentação, tais sistemas operam sem intenção humana, carecem de capacidade de formular uma motivação jurídica propriamente dita e estão sujeitos a interpretações errôneas, redutivas ou marginalizadas do direito e à utilização de fontes inadequadas. O desafio agora não é apenas modelar decisões, mas tornar explícitas, analisáveis e controláveis as cadeias de razões produzidas, assegurando a qualidade, a confiabilidade e a adequada organização dos dados que as sustentam (Conselho da Europa, 2025).

O presente artigo examina esse deslocamento a partir das chamadas arquiteturas híbridas de inteligência artificial, entendidas como arranjos computacionais que combinam, de um lado, sistemas capazes de gerar respostas em linguagem natural a partir do processamento de grandes volumes de textos e, de outro, mecanismos estruturados de organização, reconstrução e avaliação de argumentos.

As abordagens híbridas, quando articulam a geração automatizada de respostas com estruturas de formalização do raciocínio jurídico, tendem a incorporar, em sua própria

concepção, exigências relativas à qualidade e à organização dos dados utilizados - como a utilização de bases estruturadas, representativas e juridicamente qualificadas -, aproximando-se, ainda, das recomendações internacionais sobre o uso de dados em sistemas de IA aplicados ao direito (Conselho da Europa, 2025).

O objetivo do trabalho consiste em investigar de que modo tais arquiteturas podem contribuir para tornar explicitáveis, isto é, passíveis de exposição, e analisáveis, isto é, suscetíveis de exame crítico, as razões que embasam as respostas produzidas por sistemas de inteligência artificial, ao permitir sua reconstrução sob a forma de cadeias argumentativas juridicamente relevantes e sua avaliação à luz de critérios como coerência, consistência e possibilidade de contestação.

Do ponto de vista metodológico, o artigo adota uma abordagem teórico-analítica baseada em revisão da literatura do campo de inteligência artificial e direito e na análise conceitual de modelos recentes de integração entre sistemas neurais e estruturas formais de argumentação. Em particular, examina-se o experimento apresentado por Yu, Liga e Markovich (2025), no qual modelos de linguagem são orientados por estruturas formais de argumentação na produção de explicações de decisões automatizadas.

A proposição desenvolvida é que o avanço dos sistemas baseados em aprendizado de máquina não elimina o papel da formalização no raciocínio jurídico computacional, mas reconfigura suas funções em três níveis: como estrutura constitutiva do raciocínio decisório, como mecanismo de orientação da geração textual e como instrumento de reconstrução e avaliação da coerência justificatória das respostas produzidas.

Em sua dimensão mais sólida, essa última função é mais que um mecanismo técnico de análise. Ao permitir a reconstrução das razões subjacentes às respostas e sua submissão a critérios jurídicos de justificação, ela se aproxima de uma condição de legitimidade das decisões assistidas por inteligência artificial nas práticas jurídicas.

Conclui-se, assim, que a inteligência artificial aplicada ao direito não automatiza a interpretação jurídica em sentido forte. O que se observa, antes, é um processo de procedimentalização da produção e do controle das justificações jurídicas, no qual estruturas formais passam a funcionar como instrumentos de organização, avaliação e legitimação das razões apresentadas por sistemas automatizados.

2. A Formalização do Raciocínio Jurídico na IA Simbólica

Desde as décadas de 1970 e 1980, o campo da inteligência artificial e direito desenvolveu uma extensa agenda de pesquisa voltada à formalização computacional do raciocínio jurídico. O pressuposto que orientava essa tradição consistia na ideia de que, embora o direito envolva conflitos normativos, ambiguidades interpretativas e decisões baseadas em princípios, sua estrutura argumentativa poderia ser representada por meio de modelos formais capazes de explicitar as relações entre normas, exceções e argumentos jurídicos.

Os primeiros sistemas especialistas jurídicos buscaram representar a aplicação do direito por meio de regras condicionais e mecanismos de inferência lógica, partindo da hipótese de que a decisão jurídica poderia ser reconstruída como uma sequência de inferências baseadas em conhecimento jurídico previamente formalizado e implementado no sistema.

Tornou-se rapidamente evidente, contudo, que o direito apresenta características que desafiam abordagens puramente dedutivas. Normas jurídicas admitem exceções, podem entrar em conflito e dependem de interpretações de princípios. Por essa razão, a modelagem do raciocínio jurídico passou a explorar estruturas capazes de capturar essas características especiais.

Esse reconhecimento levou ao desenvolvimento de modelos mais sofisticados de representação do raciocínio jurídico, especialmente no campo da lógica não monotônica e da teoria da argumentação. Uma contribuição decisiva foi a proposta de frameworks abstratos de argumentação desenvolvida por Phan Minh Dung, que permitiu formalizar relações de ataque e defesa entre argumentos (Dung, 1995).

A partir dessa estrutura abstrata, pesquisadores do campo de inteligência artificial e direito desenvolveram modelos capazes de estruturar aspectos centrais da prática jurídica. Trabalhos como os de Henry Prakken e Giovanni Sartor demonstraram como estruturas argumentativas poderiam ser utilizadas para modelar conflitos normativos, incorporar exceções e reconstruir o raciocínio jurídico (Prakken; Sartor, 1996).

Com efeito, a noção de raciocínio derrotável (defeasible reasoning), segundo a qual as conclusões jurídicas possuem caráter provisório e podem ser revistas à luz de novas informações, evidencia a natureza não monotônica do direito, na medida em que a adição de premissas pode conduzir à revisão ou abandono de inferências anteriormente aceitas. Como observa Sartor, o raciocínio jurídico deve ser compreendido como um processo dialético estruturado pela interação entre argumentos concorrentes, cujas conclusões se mantêm apenas na ausência de razões contrárias (Sartor, 2005).

Nesses modelos, diferentes tipos de regras jurídicas – como regras estritas, regras derrotáveis, exceções e prioridades – são organizados em estruturas que permitem identificar qual argumento prevalece em determinado contexto interpretativo.

A ampliação valorativa do modelo foi proposta por Trevor Bench-Capon (2003), que acrescentou uma nova camada analítica ao reconhecer que muitos conflitos jurídicos refletem também tensões entre valores concorrentes. A prevalência de determinado argumento pode depender da hierarquia valorativa adotada: a segurança jurídica pode ceder diante da equidade; a eficiência pode ser limitada por garantias fundamentais. A arquitetura computacional passou, assim, a incorporar a ideia de que a racionalidade jurídica envolve escolhas principiológicas, e não apenas a aplicação de critérios formais.

Apesar dessas inflexões, a tradição simbólica de inteligência artificial jurídica compartilhou um pressuposto comum: a decisão jurídica pode ser reconstruída por meio de estruturas formais que explicitam as relações entre razões concorrentes. Nesse contexto, a formalização desempenha função constitutiva do raciocínio computacional: a decisão judicial resulta da aplicação das estruturas argumentativas previamente representadas no sistema. Essa concepção marcou grande parte da tradição simbólica da inteligência artificial jurídica.

O surgimento recente de sistemas baseados em grandes modelos de linguagem inaugura nova etapa nesse debate. Diferentemente dos sistemas simbólicos clássicos, os modelos baseados em sistemas neurais não dependem de formalização explícita de normas, regras ou princípios. Sua capacidade de produzir interpretações jurídicas plausíveis decorre de processos de aprendizagem estatística sobre grandes conjuntos de textos jurídicos. Essa mudança tecnológica altera o papel historicamente atribuído à formalização no campo da inteligência artificial jurídica.

3. A Ruptura Introduzida pelos Modelos Neurais

Os modelos neurais contemporâneos introduzem uma mudança estrutural no modo como sistemas de inteligência artificial produzem respostas jurídicas. Diferentemente dos sistemas simbólicos clássicos, esses modelos não operam mediante a aplicação de regras explicitamente representadas ou de estruturas formais de argumentação previamente definidas. Em vez disso, aprendem regularidades linguísticas presentes em grandes conjuntos de textos e passam a produzir respostas com base em padrões estatísticos associados ao contexto da pergunta formulada.

Essa capacidade de geração textual tem demonstrado resultados expressivos em tarefas relacionadas ao direito. Modelos dessa natureza podem sintetizar processos, reconstruir fundamentos normativos e formular justificações para soluções jurídicas apresentadas em linguagem natural. Em muitos casos, as respostas produzidas apresentam grau elevado de fluidez argumentativa e mobilizam categorias típicas da prática jurídica, como princípios, precedentes e interpretações sistemáticas.

Entretanto, tais respostas não resultam da aplicação de estruturas formais de raciocínio jurídico. Elas decorrem de processos estatísticos de previsão de sequências linguísticas, nos quais cada elemento textual é produzido com base na probabilidade de ocorrência de determinados padrões discursivos.

A crítica segundo a qual esses sistemas operam como "papagaios estocásticos", isto é, mecanismos que recombinam padrões linguísticos sem qualquer acesso ao significado, revela-se importante para compreender os limites epistêmicos e os riscos associados ao seu uso (Bender et al., 2021).

A aparente coerência das respostas pode mascarar a reprodução de vieses, distorções e imprecisões inerentes aos dados de treinamento, conferindo-lhes uma aparência de fundamentação que não corresponde a um processo genuíno de compreensão. Esse problema torna-se particularmente sensível em contextos institucionais como o jurídico, nos quais decisões dependem de elevados padrões de confiabilidade.

Em regiões marcadas por desigualdades estruturais na produção e no acesso a dados, como na América Latina, esses riscos são ainda mais acentuados. O primeiro relatório sobre o uso de inteligência artificial no setor de justiça aponta que países não europeus vêm adotando, de forma crescente, sistemas de inteligência artificial generativa em infraestruturas e serviços de empresas estrangeiras, especialmente norte-americanas, com forte dependência de computação em nuvem e limitada capacidade local de desenvolvimento de soluções próprias, em razão de restrições técnicas e econômicas (CEPEJ, 2025).

Nesse ponto, torna-se necessário distinguir dois planos na avaliação dessas tecnologias: o plano da plausibilidade discursiva e o plano da adequação da fundamentação jurídica.

A plausibilidade discursiva refere-se à capacidade de produzir respostas linguisticamente coerentes e aparentemente consistentes com o contexto jurídico da questão apresentada. Nesse sentido, modelos de linguagem tendem a gerar enunciados e textos fluentemente estruturados, ainda que desvinculados de qualquer compromisso com compreensão semântica ou com a verdade das proposições enunciadas (Bender et al., 2021).

Já a adequação da fundamentação jurídica diz respeito à possibilidade de reconstruir a decisão como resultado de um encadeamento de razões juridicamente relevantes, estruturado de modo a permitir sua avaliação crítica à luz de normas, princípios e critérios institucionais de decisão. Essa dimensão pressupõe a explicitação das premissas normativas mobilizadas e dos critérios de ponderação e resolução de conflitos entre razões concorrentes, o que não pode ser reduzido à mera coerência formal do discurso.

Isso acontece porque o raciocínio jurídico pode ser compreendido como um processo estruturado que envolve, de um lado, a construção de teorias normativas, compreendendo fatores, regras, valores e suas relações de dependência e prioridade e, de outro, a sua aplicação a casos concretos, em um movimento contínuo de abstração e concretização (Sartor, 2006; Bench-Capon; Sartor, 2003). Como destaca Sartor, a fundamentação jurídica exige a articulação explícita de razões e critérios decisórios passíveis de reconstrução e controle crítico.

A dissociação entre plausibilidade discursiva e fundamentação jurídica adequada possui implicações diretas para o uso institucional da inteligência artificial no campo jurisdicional, na medida em que evidências empíricas demonstram que modelos neurais podem alcançar elevado desempenho em tarefas envolvendo argumentação sem realizar apreensão do significado dos argumentos, operando antes por meio de correlações superficiais presentes nos dados de treinamento (Niven; Kao, 2019).

A consequência desse fenômeno é um deslocamento significativo no modo como o problema da decisão automatizada deve ser compreendido no campo da inteligência artificial e direito. Nos modelos simbólicos clássicos, a formalização do raciocínio jurídico precedia a decisão: a conclusão resultava da aplicação de estruturas argumentativas previamente formalizadas. Nos sistemas neurais contemporâneos, ao contrário, a resposta textual pode anteceder qualquer explicitação formal de sua estrutura justificatória.

Esse movimento altera o papel historicamente atribuído à formalização na inteligência artificial jurídica. Se, na tradição simbólica, ela desempenhava função constitutiva da decisão, nos sistemas neurais deixa de ocupar essa posição. O problema central passa, assim, da modelagem formal para a explicitação e o controle das fundamentações produzidas.

É nesse ponto que se torna relevante examinar o desenvolvimento de arquiteturas híbridas, que combinam modelos neurais de geração textual com estruturas formais de análise argumentativa.

4. A Reconfiguração Funcional da Formalização nas Arquiteturas Híbridas

A transformação introduzida pelos modelos neurais não implica o desaparecimento da formalização do raciocínio jurídico no campo da inteligência artificial aplicada ao direito. O que se propõe é a reconfiguração do papel desempenhado pelas estruturas formais no processo de produção e avaliação das decisões assistidas por inteligência artificial. Se, na tradição simbólica, a formalização constituía o próprio núcleo do processo decisório, nas arquiteturas contemporâneas ela passa a desempenhar funções distintas.

Essas arquiteturas são descritas como sistemas híbridos. Trata-se de arranjos computacionais que articulam dois tipos distintos de componentes: de um lado, modelos de aprendizado de máquina capazes de produzir respostas em linguagem natural com base em padrões estatísticos; de outro, estruturas formais destinadas a explicitar, reconstruir ou avaliar a coerência justificatória dessas respostas.

A investigação sobre tais arquiteturas explora a complementaridade entre essas duas dimensões. Os modelos neurais oferecem flexibilidade interpretativa e capacidade de lidar com contextos linguísticos complexos, enquanto as estruturas formais proporcionam transparência estrutural e capacidade de explicitação das razões mobilizadas. Nesse contexto, a formalização pode desempenhar três funções principais no funcionamento desses sistemas.

A primeira corresponde à função constitutiva, característica da tradição simbólica da inteligência artificial jurídica. Nessa configuração, a estrutura formal do raciocínio precede a decisão e determina o modo pelo qual argumentos são avaliados.

A segunda função pode ser descrita como orientadora da geração. Em determinadas arquiteturas híbridas, estruturas argumentativas são utilizadas para disciplinar o próprio processo de geração das respostas dadas por modelos neurais. Nesse caso, a formalização não atua apenas como mecanismo de análise posterior, mas influencia diretamente a forma como a justificativa é produzida. Essa orientação pode ocorrer por meio da incorporação de esquemas argumentativos, estruturas dialéticas ou instruções que induzem o modelo a organizar a resposta segundo determinados padrões de argumentação.

Resultados empíricos recentes ilustram esse tipo de abordagem. Em estudo apresentado na International Conference on Artificial Intelligence and Law (ICAIL 2025), Yu, Liga e Markovich investigam como estruturas simbólicas de argumentação podem ser utilizadas para orientar o comportamento de grandes modelos de linguagem em contextos de explicação de decisões automatizadas.

No experimento desenvolvido pelos autores, um modelo de linguagem é instruído, por meio de técnicas de prompt engineering, a produzir respostas estruturadas segundo um esquema argumentativo explícito. Em vez de gerar explicações livres em linguagem natural, o modelo é

induzido a organizar a resposta em etapas que identificam a decisão proposta, explicitam as premissas normativas relevantes, apresentam possíveis objeções ou contra-argumentos e indicam razões pelas quais determinadas interpretações devem prevalecer.

Os autores aplicam essa metodologia a um caso prototípico de decisão automatizada envolvendo a concessão de um empréstimo bancário, no qual diferentes razões pró e contra a operação entram em conflito. O cenário é formalizado como um sistema de argumentação abstrata, em que cada argumento - como a confiabilidade do cliente, os riscos à liquidez do banco ou o impacto ambiental da atividade financiada - é representado em uma rede de ataques e contra-ataques. A partir dessa modelagem, o modelo é orientado a reconstruir a decisão por meio da aplicação de semântica fundamentada (grounded semantics), que determina quais argumentos devem ser aceitos, rejeitados ou permanecem indeterminados, e da simulação de jogos de discussão (discussion games), nos quais posições opostas são testadas iterativamente até a estabilização de uma conclusão.

No caso analisado, a decisão de negar o empréstimo decorre da prevalência de argumentos relacionados a riscos ambientais e à fragilidade da confiabilidade do cliente, ao passo que outros elementos, como a liquidez do banco, permanecem irrelevantes por não influenciarem o resultado sob essa estrutura. Esse arranjo permite não apenas explicar por que a decisão foi tomada, mas também indicar quais premissas poderiam ser contestadas para alterar o resultado, articulando, assim, o direito à explicação com o direito à contestação.

Ao estruturar a geração textual segundo padrões argumentativos reconhecíveis, o sistema passa a produzir explicações mais adequadas a contextos institucionais nos quais os indivíduos devem ser capazes de compreender, questionar e contestar decisões automatizadas (Yu; Liga; Markovich, 2025).

Percebe-se, assim, que as arquiteturas híbridas têm a capacidade de refletir a própria natureza do direito, sendo capazes de capturar a não monotonicidade característica do raciocínio jurídico e de permitir a revisão de conclusões, por representar adequadamente conflitos normativos por meio da estrutura argumento versus contra-argumento.

Para além dessa dimensão, sustenta-se, neste trabalho - o que constitui a sua principal contribuição -, a existência de uma terceira função consistente na possibilidade de explicabilidade *ex post* das decisões produzidas por esses sistemas. Trata-se da capacidade de reconstruir, a partir da resposta gerada, a cadeia de razões subjacente, tornando-a suscetível de exame, crítica e controle à luz de modelos formais de argumentação jurídica.

Essa terceira função corresponde, portanto, à regulação da coerência justificatória das respostas produzidas por sistemas híbridos. Ainda que a geração da resposta ocorra no interior

de modelos de linguagem, estruturas formais de argumentação podem ser mobilizadas para reconstruir e examinar a cadeia de razões implicitamente subjacente à resposta, que passa, assim, a constituir objeto de reconstrução argumentativa. Isso permite avaliar a consistência e a adequação das justificações apresentadas à luz de modelos formais de raciocínio jurídico.

Pode-se objetar que a reconstrução argumentativa *ex post* não corresponde necessariamente às operações internas dos modelos de linguagem, de modo que as razões reconstruídas não seriam as "reais" razões da decisão. Contudo, a função dessas estruturas não consiste em descrever causalmente o processo interno do modelo, mas em tornar a resposta suscetível de avaliação à luz de critérios jurídicos. Nesse sentido, a reconstrução argumentativa não visa a reproduzir o funcionamento do sistema, mas a viabilizar sua inserção em práticas institucionais de justificação.

Tal possibilidade se sustenta na própria caracterização do raciocínio jurídico como não monotônico e sensível a exceções, o que exige que explicações não se limitem à apresentação de razões favoráveis, mas incorporem também contra-argumentos e potenciais derrotadores. Nessa perspectiva, a explicação deixa de ser meramente descritiva e se torna estruturalmente aberta à revisão e à impugnação, aproximando-se de uma função de controle jurídico das decisões automatizadas (Prajescu; Confalonieri, 2025).

Nessa linha, a explicabilidade não deve ser compreendida como simples transparência do funcionamento interno do modelo, mas como uma forma de justificação situada, orientada à avaliação por atores jurídicos e à sua inserção nos mecanismos institucionais de revisão e controle decisório.

Ademais, a literatura recente tem evidenciado um descompasso estrutural entre as formas de explicação oferecidas por sistemas de inteligência artificial e as exigências próprias do raciocínio jurídico. Enquanto os sistemas tendem a produzir explicações de natureza técnica, frequentemente genéricas e desvinculadas do contexto decisório, o direito demanda justificações situadas, normativamente orientadas e sensíveis às especificidades do caso concreto (Richmond et al., 2024).

Nesse contexto, tem sido proposta a adoção de um pluralismo estruturado da explicabilidade, segundo o qual não há uma única forma adequada de explicar decisões jurídicas, mas sim diferentes modalidades explicativas, dependentes do tipo de decisão, do regime probatório e das exigências institucionais envolvidas (Richmond et al., 2024).

Essa perspectiva reforça a necessidade de compreender a explicação não apenas como um atributo técnico dos sistemas, mas como um elemento integrado às práticas jurídicas. É nesse ponto que se insere a terceira função aqui defendida: ao permitir a reconstrução

argumentativa das razões subjacentes às respostas produzidas, cria-se uma via para reconfigurar explicações técnicas em justificações juridicamente relevantes, aptas a serem avaliadas, criticadas e eventualmente revistas.

A coexistência dessas três funções revela que a formalização deixa de ocupar posição única no processo decisório. Em arquiteturas híbridas, estruturas formais podem atuar em diferentes momentos do processo: antes da geração da resposta, orientando o raciocínio; durante a geração, disciplinando a organização da justificativa; ou após a produção da resposta, reconstruindo e avaliando a coerência das razões apresentadas.

Essa reconfiguração aproxima o debate técnico sobre arquiteturas híbridas de discussões jurídicas mais amplas acerca da exigência de explicação em decisões públicas. No direito administrativo europeu, o dever de fundamentação constitui um mecanismo central de controle do poder público. Como observam Melanie Fink e Michèle Finck, a obrigação de indicar as razões que sustentam uma decisão administrativa permite que os indivíduos compreendam o fundamento da medida que os afeta e constitui condição indispensável tanto para o exercício da revisão judicial quanto para a possibilidade de contestação das decisões administrativas (Fink; Finck, 2022).

É nesse horizonte institucional que se insere o debate contemporâneo sobre explicabilidade em sistemas de decisão assistidos por inteligência artificial. Como observa Davide Liga, a exigência de explicação nesses contextos não deve ser compreendida apenas como um problema técnico de transparência algorítmica, mas como uma condição para que as decisões produzidas com o auxílio de sistemas automatizados possam ser compreendidas, justificadas e contestadas no interior das instituições jurídicas (Liga, 2024). A explicação desempenha, assim, papel essencial na garantia de mecanismos de responsabilização, revisão e contestação das decisões dadas por sistemas automatizados.

Sob essa perspectiva, arquiteturas híbridas podem desempenhar papel institucional relevante. Ao possibilitar a explicitação de premissas, a reconstrução de cadeias argumentativas e a identificação de conflitos normativos, tais sistemas contribuem para tornar analisáveis as justificações produzidas por modelos neurais.

Em vez de automatizar a interpretação jurídica em sentido forte, essas arquiteturas tendem a estruturar procedimentos de produção, explicitação e controle das justificações jurídicas, alinhando o seu funcionamento às exigências institucionais associadas ao dever de fundamentação e ao direito à explicação. Essa função assume, assim, a forma de condição de legitimidade das decisões assistidas por inteligência artificial.

A necessidade de integrar estruturas formais ao funcionamento de modelos neurais torna-se ainda mais evidente quando examinada à luz das outras pesquisas recentes apresentadas na ICAIL 2025, que revelam um movimento convergente no sentido de articular modelos neurais a estruturas formais de argumentação como forma de enfrentar os limites epistêmicos da inteligência artificial generativa no domínio jurídico.

Nesse cenário, Gray, Savelka, Oliver e Ashley demonstram como grandes modelos de linguagem podem ser mobilizados para identificar não apenas a presença de fatores fáticos e normativos, mas também a sua magnitude em casos concretos. Nessa arquitetura, o modelo neural atua como um extrator semântico cujos resultados alimentam uma estrutura formal de argumentação, garantindo que a saída seja juridicamente inteligível e argumentativamente rigorosa (Gray et al., 2025).

A urgência dessa disciplina formal é corroborada, também, por investigações sobre a fragilidade do raciocínio interno dos modelos generativos. Em linha convergente, Steging, Renooij e Verheij propõem tarefas parametrizadas de raciocínio argumentativo para avaliar o desempenho de grandes modelos de linguagem. Ao submeter esses modelos a grafos de ataque argumentativo de complexidade variável, os autores constatam que os sistemas de estado da arte frequentemente falham em resolver problemas lógicos, mesmo em níveis baixos de complexidade (Steging; Renooij; Verheij, 2025). Tais resultados indicam que o comportamento inferencial dos modelos generativos permanece frágil e opaco, o que inviabiliza sua aplicação responsável em domínios sensíveis como o direito sem a mediação de estruturas simbólicas de controle.

É nesse ponto de intersecção entre a opacidade estatística e a exigência de justificação que a argumentação computacional se consolida como o mecanismo privilegiado de explicabilidade jurídica. Essa compreensão é reforçada por Rotolo e Sartor, para quem a explicação no direito não pode ser reduzida a uma descrição causal do funcionamento interno de um algoritmo. Ela deve constituir uma justificação formal que conecte premissas normativas e fáticas a uma conclusão, tornando-a suscetível de contestação. Os autores demonstram que a justificação formal e a explicação estão conceitualmente interligadas, e que a noção de explicação normativa estável pode ser modelada em lógicas derrotáveis (Rotolo; Sartor, 2023).

Essa perspectiva é operacionalizada em arquiteturas que utilizam a argumentação para justificar predições de modelos de caixa-preta. Peters, Bex e Prakken propõem um sistema que emprega argumentação baseada em casos para explicar as saídas de classificadores de aprendizado de máquina. Ao argumentar sobre as diferenças e semelhanças entre o caso atual

e os precedentes, o sistema utiliza o conhecimento de domínio como restrições sobre os argumentos admissíveis (Peters; Bex; Prakken, 2025).

Essas abordagens materializam a terceira função da formalização, correspondente à regulação da coerência justificatória, permitindo que uma decisão gerada por um modelo opaco seja reconstruída e avaliada segundo os cânones do raciocínio jurídico.

A convergência entre essas diferentes linhas de pesquisa aponta para o que Floris Bex descreve como a necessidade de um ecossistema transdisciplinar para o futuro da inteligência artificial no direito. Em seu discurso presidencial na ICAIL, Bex argumenta que o avanço responsável da área depende da combinação entre a pesquisa em sistemas orientados a dados, como a inteligência artificial generativa, e a tradição da inteligência artificial baseada em conhecimento (Bex, 2024). Essa síntese neuro-simbólica não apenas potencializa a capacidade de processamento de linguagem natural, mas assegura que as aplicações jurídicas permaneçam ancoradas em modelos formais de argumentação, prova e justificação.

Em suma, a literatura contemporânea consolida o entendimento de que a explicabilidade em sistemas de decisão assistidos por inteligência artificial transcende o problema da transparência algorítmica, configurando-se como um requisito normativo de justificação e contestabilidade. As arquiteturas híbridas, ao articularem a expressividade dos modelos neurais com o rigor das estruturas simbólicas, oferecem o arranjo tecnológico mais promissor para alinhar a automação decisória às exigências institucionais do Estado de Direito. Nesse novo paradigma, a formalização não desaparece, mas se reconfigura como a gramática essencial que traduz predições estatísticas em razões jurídicas legítimas.

5. Considerações Finais

O desenvolvimento recente de grandes modelos de linguagem introduziu uma transformação relevante no campo da inteligência artificial aplicada ao direito. Enquanto as abordagens clássicas buscavam formalizar diretamente o raciocínio responsável pela decisão, os sistemas contemporâneos são capazes de produzir respostas juridicamente plausíveis por meio de processos de geração linguística estatística que não dependem de representação normativa explícita. Essa mudança desloca o problema central da área: em vez de se concentrar exclusivamente na modelagem formal da decisão, a investigação passa a enfrentar o desafio de tornar analisáveis, controláveis e institucionalmente integráveis as justificações produzidas por sistemas neurais.

Este artigo argumentou que arquiteturas híbridas constituem um caminho promissor para lidar com esse deslocamento. Ao combinar modelos neurais de geração textual com estruturas formais de reconstrução e análise argumentativa, esses sistemas permitem reconfigurar o papel tradicionalmente atribuído à formalização no raciocínio jurídico computacional. A formalização deixa de operar apenas como estrutura constitutiva da decisão e passa a desempenhar funções múltiplas: pode estruturar previamente o raciocínio decisório, orientar a produção das justificações ou servir como instrumento de reconstrução e avaliação da adequação jurídica das respostas produzidas.

Essa reconfiguração não é apenas uma adaptação técnica, mas revela uma mudança mais profunda na forma como se compreende a racionalidade jurídica em contextos mediados por inteligência artificial. Se, na tradição simbólica, a racionalidade do sistema dependia da fidelidade entre a formalização e o processo decisório, nas arquiteturas contemporâneas a racionalidade passa a depender da possibilidade de reconstrução, escrutínio e contestação das razões apresentadas. Em outras palavras, desloca-se o foco da produção da decisão para a sua justificabilidade.

Nesse contexto, a terceira função da formalização, entendida como mecanismo de reconstrução e avaliação da coerência justificatória, assume novo papel. Ao possibilitar a submissão das decisões assistidas por inteligência artificial a critérios jurídicos de justificação, controle e contestação, essa função complementa as demais e redefine o próprio sentido da explicabilidade no direito. A explicação é compreendida como uma prática argumentativa orientada à avaliação institucional.

Essa estruturação tem implicações para a teoria do direito e para a prática jurisdicional. Em primeiro lugar, reforça a ideia de que a legitimidade das decisões não decorre da origem, humana ou automatizada, mas da qualidade das razões que podem ser publicamente reconstruídas, examinadas e eventualmente rejeitadas. Em segundo lugar, evidencia que o uso de sistemas de inteligência artificial no direito exige a criação de procedimentos capazes de inserir essas decisões em circuitos institucionais de controle, nos quais a explicação desempenha função estruturante.

Além disso, a análise desenvolvida permite identificar que o desafio contemporâneo não se limita à opacidade dos modelos, mas envolve também a qualidade, a curadoria e a governança dos dados que alimentam esses sistemas. A plausibilidade discursiva das respostas pode ocultar distorções derivadas de bases de dados enviesadas, incompletas ou descontextualizadas, o que reforça a necessidade de mecanismos complementares de validação, auditoria e supervisão. Nesse sentido, arquiteturas híbridas não apenas contribuem para a

explicitação das razões, mas tendem a incorporar, em sua própria concepção, exigências relativas à qualidade e à organização dos insumos informacionais que sustentam as decisões.

Outro aspecto relevante diz respeito à relação entre explicabilidade e contestabilidade. Ao permitir a reconstrução argumentativa das decisões, tais arquiteturas tornam as razões visíveis e indicam os pontos nos quais essas razões podem ser questionadas, revistas ou superadas. A explicação abrange a descrição do porquê de uma decisão e se projeta como condição de possibilidade para o exercício do contraditório e para a efetividade do controle jurídico.

Sob essa perspectiva, a inteligência artificial aplicada ao direito não automatiza a interpretação jurídica em sentido forte. O que se observa é o advento de um modelo no qual a produção e o controle das justificações passam a ser estruturados por procedimentos tecnicamente mediados. Trata-se, portanto, de um processo de procedimentalização da racionalidade jurídica, no qual a validade prática das decisões depende de sua capacidade de se inserir em estruturas formais de explicitação, avaliação e crítica.

Essa conclusão permite também delimitar os limites dessas tecnologias. Arquiteturas híbridas não eliminam a necessidade de interpretação jurídica institucional, nem substituem o papel dos atores jurídicos na atribuição de significado, na ponderação de valores e na definição de critérios decisórios. Ao contrário, reforçam a centralidade dessas atividades, ao exigir que as justificações produzidas por sistemas automatizados sejam continuamente reconstruídas e avaliadas à luz de parâmetros jurídicos.

Em síntese, a principal contribuição deste artigo consiste em demonstrar que o impacto dos modelos neurais no direito não reside na substituição da racionalidade jurídica, mas na transformação das condições sob as quais ela se manifesta. A formalização, longe de desaparecer, assume novas funções e passa a operar como instrumento de mediação entre sistemas automatizados e práticas institucionais de justificação. É nesse espaço que se delineia o futuro da inteligência artificial no direito.

Referências

BENCH-CAPON, Trevor. Persuasion in practical argument using value-based argumentation frameworks. *Journal of Logic and Computation*, Oxford, v. 13, n. 3, p. 429-448, 2003. DOI: https://doi.org/10.1093/logcom/13.3.429.

BENDER, Emily M.; GEBRU, Timnit; McMILLAN-MAJOR, Angelina; SHMITCHELL, Shmargaret. On the dangers of stochastic parrots: can language models be too big? In: ACM

CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY (FAccT '21), 2021. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. New York: ACM, 2021. DOI: https://doi.org/10.1145/3442188.3445922.

BEX, Floris. AI, law and beyond: a transdisciplinary ecosystem for the future of AI & law. *Artificial Intelligence and Law*, 2024. DOI: https://doi.org/10.1007/s10506-024-09404-y.

CONSELHO DA EUROPA. Comissão Europeia para a Eficiência da Justiça (CEPEJ). Artificial Intelligence Advisory Board (AIAB). Reflections of the AIAB on the use of artificial intelligence in judicial systems. Strasbourg: Council of Europe, 2025. Disponível em: https://rm.coe.int/cepej-aiab-2025-1rev5-en-reflections-of-the-aiab-on-the-use-of-artific/1680b42a56. Acesso em: 20 mar. 2026.

CONSELHO DA EUROPA. Comissão Europeia para a Eficiência da Justiça (CEPEJ). Advisory Board on Artificial Intelligence (AIAB). 1st AIAB report on the use of artificial intelligence (AI) in the judiciary based on the information contained in the resource centre on cyberjustice and AI. Strasbourg: Council of Europe, 2025. (CEPEJ-AIAB(2024)4Rev5). Disponível em: https://rm.coe.int/cepej-aiab-2024-4rev5-en-first-aiab-report-2788-0938-9324-v-1/1680b49def. Acesso em: 20 mar. 2026.

DUNG, Phan Minh. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, v. 77, n. 2, p. 321-357, 1995. DOI: https://doi.org/10.1016/0004-3702(94)00041-X.

FINK, Melanie; FINCK, Michèle. Reasoned A(I)dministration: explanation requirements in EU law and the automation of public administration. *European Law Review*, v. 47, n. 3, p. 376-392, 2022. Disponível em: https://hdl.handle.net/1887/3439725. Acesso em: 20 mar. 2026.

GRAY, Morgan; SAVELKA, Jaromír; OLIVER, Will; ASHLEY, Kevin D. Generating legal arguments with automatically identified factor magnitudes. In: INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW (ICAIL), 20., 2025. *Proceedings of the 20th International Conference on Artificial Intelligence and Law*. New York: ACM, 2025. DOI: https://doi.org/10.1145/3769126.3769212.

LIGA, Davide. The interplay between lawfulness and explainability in automated decision-making of EU administration. In: HOFMANN, Herwig C. H.; PFLÜCKE, Felix (org.). *Governance of automated decision-making and EU law*. Oxford: Oxford University Press, 2024. DOI: https://doi.org/10.1093/9780198919575.003.0009.

NIVEN, Timothy; KAO, Hung-Yu. Probing neural network comprehension of natural language arguments. In: ANNUAL MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS (ACL), 57., 2019, Florence, Italy. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Florence: Association for Computational Linguistics, 2019. p. 4658-4664. DOI: https://doi.org/10.18653/v1/P19-1459.

PETERS, Jeroen; BEX, Floris; PRAKKEN, Henry. Justifying black-box predictions with domain knowledge. In: INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW (ICAIL), 20., 2025. *Proceedings of the 20th International Conference on Artificial Intelligence and Law*. New York: ACM, 2025. DOI: https://doi.org/10.1145/3769126.3769253.

PRAJESCU, Andrada Iulia; CONFALONIERI, Roberto. Argumentation-based explainability for legal AI: comparative and regulatory perspectives. arXiv, 2025. Disponível em: https://arxiv.org/abs/2510.11079. Acesso em: 20 mar. 2026. DOI: https://doi.org/10.48550/arXiv.2510.11079.

PRAKKEN, Henry; SARTOR, Giovanni. A dialectical model of assessing conflicting arguments in legal reasoning. *Artificial Intelligence and Law*, Dordrecht, v. 4, n. 3-4, p. 331-368, 1996. Disponível em: https://webspace.science.uu.nl/~prakk101/pubs/DefArgAILaw96.pdf. Acesso em: 20 mar. 2026.

RICHMOND, Karen McGregor; MUDDAMSETTY, Satya; GAMMELTOFT-HANSEN, Thomas; OLSEN, Henrik Palmer. Explainable AI and law: an evidential survey. 2024. Disponível em: https://vbn.aau.dk/en/publications/explainable-ai-and-law-an-evidential-survey/. Acesso em: 16 mar. 2026.

ROTOLO, Antonino; SARTOR, Giovanni. Argumentation and explanation in the law. *Frontiers in Artificial Intelligence*, v. 6, e1130559, 2023. DOI: https://doi.org/10.3389/frai.2023.1130559.

SARTOR, Giovanni. Legal reasoning: a cognitive approach to the law. Dordrecht: Springer, 2005.

SARTOR, Giovanni. Syllogism and defeasibility: a comment on Neil MacCormick's rhetoric and the rule of law. *Northern Ireland Legal Quarterly*, v. 59, p. 20-31, 2006.

STEGING, Cees; RENOOIJ, Silja; VERHEIJ, Bart. Parameterized argumentation-based reasoning tasks for benchmarking generative language models. In: INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW (ICAIL), 20., 2025. *Proceedings of the 20th International Conference on Artificial Intelligence and Law*. New York: ACM, 2025. DOI: https://doi.org/10.1145/3769126.3769230.

YU, Liuwen; LIGA, Davide; MARKOVICH, Réka. Addressing the right to explanation and the right to challenge through hybrid-AI: symbolic constraints over large language models via prompt engineering. In: INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW (ICAIL), 20., 2025, Chicago, IL. *Proceedings of the 20th International Conference on Artificial Intelligence and Law*. New York: ACM, 2025. p. 239–247. DOI: https://doi.org/10.1145/3769126.3769252.