

## **INTELIGÊNCIA ARTIFICIAL EXPLICÁVEL NO PODER JUDICIÁRIO BRASILEIRO: análise no Tribunal de Justiça de Rondônia**

## **EXPLAINABLE ARTIFICIAL INTELLIGENCE IN THE BRAZILIAN JUDICIARY: analysis in the Court of Justice of Rondônia**

**Eduardo Abílio Kerber Diniz<sup>1</sup>**

**Alan Almeida do Amaral<sup>2</sup>**

### **RESUMO:**

Este artigo analisa o uso da inteligência artificial (IA) no Poder Judiciário brasileiro, com foco no desenvolvimento e aplicação de sistemas de IA explicáveis no Tribunal de Justiça de Rondônia (TJRO). Objetiva, mediante pesquisa qualitativa, discutir a importância da transparência e da compreensibilidade das decisões automatizadas para garantir a justiça e a confiança pública. Como resultado, conclui que o TJRO demonstra o potencial dessas tecnologias para reformular os processos judiciais e se preocupa com a necessidade de implementar práticas responsáveis e éticas que assegurem a explicabilidade e a transparência da IA no Poder Judiciário.

**PALAVRAS-CHAVE:** Inteligência Artificial. Poder Judiciário. Tribunal de Justiça de Rondônia. Transparência. Decisões Automatizadas.

### **ABSTRACT:**

This article examines the use of artificial intelligence (AI) in the Brazilian Judiciary, focusing on the development and application of explainable AI systems within the Court of Justice of Rondônia (TJRO). Through qualitative research, it aims to discuss the importance of transparency

<sup>1</sup> Juiz Substituto do Tribunal de Justiça de Rondônia. Professor Adjunto da Universidade Federal de Rondônia. Mestre e Doutorando em Função Social do Direito pela Faculdade Autônoma de Direito. Graduação em Direito pelo Centro Universitário de Brasília. E-mail: [eduardoabilio@tjro.jus.br](mailto:eduardoabilio@tjro.jus.br)

<sup>2</sup> Advogado. Pós-Graduando em Direito Regulatório e Concorrencial. Graduação em Direito pela Universidade Federal de Rondônia. E-mail: [alan.amaral@roumieamaral.com.br](mailto:alan.amaral@roumieamaral.com.br)

and comprehensibility in automated decisions to ensure justice and public trust. As a result, it concludes that the TJRO demonstrates the potential of these technologies to reshape judicial processes and highlights the need to implement responsible and ethical practices that ensure AI explainability and transparency within the Judiciary.

**KEYWORDS:** Artificial Intelligence. Judiciary. Court of Justice of Rondônia. Transparency. Automated Decisions.

## 1. INTRODUÇÃO

A rápida evolução das tecnologias de inteligência artificial (IA) tem transformado diversos setores ao redor do mundo, e o sistema judiciário não é exceção. Nos últimos anos, a adoção de ferramentas de IA nos tribunais tem sido impulsionada pela promessa de aumentar a eficiência, reduzir custos e melhorar a acessibilidade dos serviços judiciais. No entanto, essa transformação tecnológica traz consigo desafios significativos, especialmente em relação à transparência e à explicabilidade das decisões geradas ou assistidas por sistemas de IA.

Atento à relevância do tema, o Conselho Nacional de Justiça (CNJ) aprovou a Resolução n. 615, de 11 de março de 2025, que estabelece diretrizes para assegurar o desenvolvimento, utilização e governança de soluções desenvolvidas com recursos de inteligência artificial no Poder Judiciário. Ela é a sucessora da Resolução n. 332/2020, que disciplinava o tema de forma menos abrangente e desenhou os primeiros parâmetros sobre o uso de IA nos tribunais brasileiros. Na norma atual, o CNJ propõe um arranjo orientado pela transparência e pela explicabilidade (art. 3º, inciso II), cuja aferição ocorrerá por indicadores públicos e informacionalmente acessíveis.

Notadamente, a inteligência artificial explicável (XAI, do inglês *Explainable Artificial Intelligence*) emerge como uma abordagem crucial nesse contexto, pois busca garantir que as decisões automatizadas permaneçam compreensíveis e auditáveis tanto para os operadores do direito quanto para as partes envolvidas. A transparência é fundamental para manter a confiança pública no sistema judicial,

assegurando que as decisões sejam justas e baseadas em processos claros e inteligíveis. Sem essa clareza, existe o risco de minar a legitimidade das decisões judiciais, especialmente aquelas que impactam diretamente a vida das pessoas.

O Tribunal de Justiça de Rondônia (TJRO) se destaca como um exemplo notável de inovação no uso de tecnologias judiciais no Brasil. A escolha do TJRO como estudo de caso para a análise da implementação da IA explicável no sistema judicial é justificada pela sua postura proativa na adoção de soluções tecnológicas avançadas. O tribunal tem liderado iniciativas que buscam integrar sistemas de IA de maneira responsável e ética, destacando-se em um cenário nacional como referência em inovação judiciária.

Ao explorar como o TJRO implementa tecnologias de IA, este artigo busca oferecer *insights* sobre as práticas promissoras e os desafios enfrentados por um tribunal que está na vanguarda da transformação digital. A análise do caso de Rondônia não apenas ilustra o potencial da IA para melhorar os processos judiciais, mas também esclarece a importância da explicabilidade na salvaguarda dos princípios de justiça e equidade no ambiente jurídico.

## 2. O PAPEL DA INTELIGÊNCIA ARTIFICIAL NO PODER JUDICIÁRIO

### 2.1 Histórico e Evolução da Inteligência Artificial no Sistema Judicial Brasileiro

O uso da inteligência artificial (IA) no Poder Judiciário brasileiro tem suas raízes nas décadas recentes, impulsionado pela busca incessante por eficiência e modernização do sistema judicial. A evolução tecnológica no Brasil começou a ganhar tração nos anos 2000, com iniciativas voltadas para a digitalização de processos e a implementação de sistemas eletrônicos de tramitação processual. Esses sistemas serviram de base para a introdução gradual de ferramentas mais sofisticadas de IA.

Nos últimos anos, o Conselho Nacional de Justiça (CNJ) e diversos tribunais estaduais, incluindo o Tribunal de Justiça de Rondônia (TJRO), têm dedicado esforços significativos à incorporação de soluções de IA. Estas soluções abrangem desde a automação de tarefas administrativas, como o agendamento de audiências e a gestão de documentos, até a análise de dados complexos para auxiliar na tomada de decisões judiciais.

Um marco importante neste contexto foi a implementação de sistemas de IA para triagem processual e análise preditiva. Esses sistemas são capazes de identificar padrões em grandes volumes de dados jurídicos, permitindo maior celeridade na tramitação dos processos e contribuindo para a diminuição do acúmulo de casos pendentes. Além disso, iniciativas como a criação de assistentes virtuais e o uso de algoritmos de *machine learning* para suporte em decisões judiciais têm ganhado destaque, ilustrando o potencial transformador da IA no sistema jurídico brasileiro.

## **2.2 Benefícios e Desafios da Inteligência Artificial no Poder Judiciário**

### **2.2.1 Benefícios**

A implementação de IA no sistema judiciário brasileiro oferece inúmeras vantagens. Primeiramente, a eficiência operacional é um dos benefícios mais notáveis. A automação de tarefas rotineiras libera o tempo dos servidores e juízes, permitindo que se concentrem em atividades mais complexas e de maior valor agregado. Isso resulta em um aumento da produtividade e na aceleração do tempo de tramitação processual, o que é crucial em um sistema frequentemente sobrecarregado.

Outro benefício significativo é a redução de custos. Com a automação de processos e a digitalização, há uma diminuição expressiva na necessidade de recursos físicos e humanos, resultando em economias substanciais para o sistema judicial.

Não se nega que a inteligência artificial pode contribuir significativamente com o acesso à ordem jurídica justa, aproximando os cidadãos do Poder Judiciário, como ocorre com a disponibilização de *chatbots* que fornecem suporte e informações aos jurisdicionados em linguagem acessível. Além disso, a IA pode contribuir para uma maior padronização das decisões judiciais, promovendo a uniformidade e a previsibilidade, elementos essenciais para a segurança jurídica.

## 2.2.2 Desafios

Em 2014, Stephen Hawking sugeriu que era necessário levar a inteligência artificial mais a sério. À época, refletia-se sobre o uso de carros autônomos e de assistentes pessoais digitais, como a Siri, o Google Now e a Cortana. O teórico listou estas discussões apenas como algo sintomático daquilo que definiu como “uma corrida armamentista de TI alimentada por investimentos sem precedentes e construída sobre uma base teórica cada vez madura”<sup>3</sup>. Sua preocupação não consistia em evitar o desenvolvimento e a implementação da IA, mas sim como aprender a evitar os riscos que lhe são intrínsecos.

Sempre há entusiasmo com a expansão das possíveis aplicações da inteligência artificial, mas a precaução devida, proposta por Hawking como alerta, não foi pautada com igual envergadura. Para ele, “todos nós deveríamos nos perguntar o que podemos fazer agora para melhorar as chances de colher os benefícios e evitar os riscos”<sup>4</sup>.

No Poder Judiciário, a leitura preditiva do Conselho Nacional de Justiça (CNJ) foi essencial à mitigação de riscos na implementação e no uso da inteligência artificial. Não por outra razão, a consideração inicial da Res. n. 332/2020 é o potencial de contribuição da IA com a agilidade e coerência no processo de tomada de decisão, mas sem dispensar o dever de atendimento de critérios éticos, de transparência, de previsibilidade, de auditabilidade e afins.

<sup>3</sup> Disponível em: <https://www.independent.co.uk/news/science/stephen-hawking-transcendence-looks-at-the-implications-of-artificial-intelligence-but-are-we-taking-ai-seriously-enough-9313474.html>. Acesso em 18 de abr. 2025.

<sup>4</sup> Idem.

A rigor, apesar dos benefícios, a integração da IA no Judiciário brasileiro enfrenta desafios consideráveis. Questões éticas e de transparência estão no centro desses desafios. A IA, especialmente quando utilizada na tomada de decisões judiciais, levanta preocupações sobre o viés algorítmico e a equidade. A falta de transparência nos processos de decisão automatizada pode comprometer a confiança pública no sistema judicial.

Além disso, há desafios relacionados à proteção de dados e à privacidade, especialmente considerando o volume de informações sensíveis processadas por sistemas judiciais. A conformidade com legislações como a Lei Geral de Proteção de Dados (LGPD) é essencial para garantir que as implementações de IA respeitem os direitos dos cidadãos.

A explicabilidade das decisões de IA é, portanto, um requisito vital para garantir que as partes envolvidas compreendam como uma determinada conclusão foi alcançada. Isso torna necessário o desenvolvimento de sistemas de IA que não apenas executem tarefas eficazmente, mas também sejam capazes de justificar suas decisões de maneira clara e comprehensível.

Em suma, enquanto a IA oferece um caminho promissor para a modernização do Judiciário brasileiro, sua implementação deve ser cuidadosamente gerida para mitigar riscos e maximizar benefícios. As lições aprendidas em tribunais como o TJRO podem servir como modelos para outras jurisdições, promovendo uma adoção responsável e eficaz da tecnologia no sistema judicial.

### **1.3 A atividade normativa do CNJ como balizadora do uso e implementação da IA no Poder Judiciário brasileiro**

A edição das Resoluções n. 332/2020 e 615/2025 pelo Conselho Nacional de Justiça bem representa o balizamento do uso e da implementação da IA no Poder Judiciário por critérios de governança,

monitoramento, uso responsável e auditabilidade. Com isso, afirmou-se o respeito aos direitos fundamentais, o incentivo ao desenvolvimento tecnológico, a segurança jurídica, a busca da eficiência, a qualidade e a uniformidade na entrega da prestação jurisdicional, além de outros valores, erigidos à condição de fundamentos e princípios.

Naquilo que foi possível antecipar, a Res. n. 332/2020 tratou do respeito aos direitos fundamentais, da não discriminação, da publicidade, da transparência, da governança, da qualidade, da segurança, do controle do usuário, das diretrizes para pesquisa, desenvolvimento e implantação de serviços de IA, da prestação de contas e da responsabilização.

Sua sucessora, a Resolução n. 615/2025, terá a função de compor uma paisagem melhor ordenada pelas fragilidades, fortalezas e limitações experimentadas no interregno entre as normas.

Dentre as modificações e inovações mais relevantes, tem-se um capítulo e um anexo dedicados à classificação de riscos, a criação do Comitê Nacional de Inteligência Artificial do Judiciário e a instituição da Avaliação de Impacto Algorítmico.

Ademais, ela estabelece o uso da IA como instrumento de apoio em determinadas ações, e não como juízo natural, mecanismos para a prevenção da discriminação e mitigação de vieses e vedações ao uso de modelos de linguagem de larga e pequena escala (*LLMs* e *SLMs*).

A utilização de inteligência artificial generativa pressupõe riscos. Alguns previstos, que demandam medidas preventivas, e outros incertos, mas possíveis, ensejando precauções. Para categorizar tais eventos, dedica-se capítulo específico do novo normativo, com a instituição do dever de avaliação das soluções que utilizem técnicas de IA para definição do seu grau de risco (art. 9º, *caput*).

As soluções de alto risco serão objeto de acompanhamento regular de auditoria e monitoramento contínuo, para supervisão e mitigação de riscos aos direitos dos jurisdicionados, à privacidade e à justiça (art. 11, §1º). O anexo da resolução trata delas com a sigla “AR”, em 05 (cinco) atividades acessórias.

Já as soluções de baixo risco devem ser monitoradas e revisadas periodicamente, justamente para que permaneçam representando baixo risco e para assegurar que as evoluções tecnológicas ou contextuais não modifiquem essa categorização (art. 11, §3º). O anexo da norma trata dessas ações com a sigla “BR”, em 08 (oito) atividades acessórias.

Há outras soluções que, categoricamente, não podem ser objeto de desenvolvimento e utilização de soluções. Elas são tidas como ações de risco excessivo (art. 10, *caput*) e, se não observadas, representam exposição à violações de segurança da informação, de direitos fundamentais e da própria independência dos magistrados. Não há prejuízo à possibilidade de que as soluções de baixo ou alto risco se convertam em risco excessivo, quando deverão ser desconstituídas, com registro no Sinapses, para análise do Comitê Nacional de Inteligência Artificial do Judiciário.

Outra novidade é o Comitê Nacional de Inteligência Artificial do Judiciário, instituído pelo art. 15, que será composto por 14 (quatorze) membros titulares e 13 (treze) suplentes, divididos por categorias e designados pelo Presidente do CNJ, sendo: dois Conselheiros do CNJ; dois juízes auxiliares e dois servidores, ambos do CNJ, com experiência na área; dois magistrados, sendo um do CJF e outro do CSJT; quatro desembargadores, sendo um representante de tribunal de justiça, um de tribunal regional federal, um de tribunal regional do trabalho e um de tribunal eleitoral; dois representantes das escolas da magistratura, sendo um da ENFAM e outro da ENAMAT; quatro magistrados a partir de indicações da AMB, ANAMATRA e AJUFE; dois representantes da OAB; dois representantes do MP; dois representantes da Defensoria Pública; e dois representantes da sociedade civil (art. 15, §1º, incisos I a X).

Em que pese a composição bem ordenada do Comitê, não se franqueou direito à voto aos representantes da OAB, do Ministério Público, da Defensoria Pública e da Sociedade Civil. No ponto, a resolução deve ser objeto de discussões. A atuação das instituições mencionadas e da sociedade civil, com direito à voto, contribuiria com o bom cumprimento das atribuições do novo órgão colegiado.

Inclusive, o Comitê, de ofício ou mediante provocação fundamentada, poderá proceder à reclassificação do grau de risco de determinada solução (baixo risco, alto risco e risco excessivo) e determinar, justificadamente, a realização da Avaliação de Impacto Algorítmico (AIA). Esta é assim conceituada pela resolução:

Análise contínua dos impactos de um sistema de IA sobre os direitos fundamentais, com a identificação de medidas preventivas, mitigadoras de danos e de maximização dos impactos positivos, sem a violação da propriedade industrial e intelectual da solução de IA utilizada.

A avaliação também deverá ser promovida, independente de determinação, pelo tribunal desenvolvedor ou contratante de solução classificada como de alto risco, observando, em todos os casos, os requisitos estabelecidos pelo Comitê e, sempre que possível, a etapa de participação social (art. 14, §1º e §2º).

Conceitualmente, ela se aproxima da Análise de Impacto Regulatório (AIR), prevista no art. 6º Lei das Agências Reguladoras Federais<sup>5</sup>. A distinção fundamental entre a AIA e a AIR é que aquela trata dos impactos de um sistema, enquanto esta se preocupa com os impactos regulatórios de atos normativos<sup>6</sup>. Sem prejuízo, as diretrizes para a tessitura da AIR poderão, quando da regulamentação da AIA, inspirar o Comitê tanto nos aspectos de incentivo à realização de consultas públicas, quanto na definição das metodologias específicas para aferição da razoabilidade do impacto algorítmico, como as análises de risco e de risco-risco, a que se referem o Decreto n. 10.411/2020, regulamentador da AIR.

5 Art. 6º, Lei das ARFs: “A adoção e as propostas de alteração de atos normativos de interesse geral dos agentes econômicos, consumidores ou usuários dos serviços prestados serão, nos termos de regulamento, precedidas da realização de Análise de Impacto Regulatório (AIR), que conterá informações e dados sobre os possíveis efeitos do ato normativo.”

6 Outra questão relevante é que a AIA se manifesta como um processo prévio de avaliação de riscos, mas também como instrumento contínuo, diferente da AIR. No âmbito regulatório, foi necessário o desenvolvimento de outro mecanismo para o acompanhamento dos impactos regulatórios. É o que se consignou como Avaliação de Resultado Regulatório (ARR), consistente em ferramenta de melhoria contínua da qualidade regulatória que permite acompanhar e avaliar o desempenho de medidas implementadas.

Esses parâmetros da resolução denotam um arranjo normativo preocupado com os direitos fundamentais dos jurisdicionados e com a concepção constitucional da justiça. O princípio do juiz natural, do devido processo legal e do contraditório efetivo foram inseridos nesse plexo. Associando-o ao uso da IA no Judiciário, desenvolveu-se o princípio do devido processo informacional. Por ele, busca-se assegurar que a utilização de inteligência artificial por instituições com poderes de tomada de decisão sobre direitos e interesses de pessoas seja guiada pela transparência, integridade e legitimidade.

Como asseveram Rodrigues, Mendonça e Nóvoa (2025, p. 4):

Isso, indubitavelmente, serve para evitar arbitrariedades nas decisões proferidas pelo Poder Judiciário, bem como reduzir as assimetrias informacionais existentes entre as tecnologias de IA e à sociedade. O princípio do devido processo informacional, portanto, é uma forma de garantir a inclusão do indivíduo no processo de tomada de decisão, permitindo-lhe se defender de maneira justa e legítima.

Essa garantia protecionista, no entanto, é vulnerabilizada quando os direitos dos jurisdicionados são expostos a sistemas de tomadas de decisão automatizadas, vez que “muitas vezes torna-se inviável o conhecimento de como determinada decisão foi tomada, ou, inclusive, se é possível contestá-la” (Rodrigues; Mendonça; Nóvoa, 2025).

Por isso, a resolução atual do CNJ busca assegurar que os recursos automatizados não substituam o juízo natural da causa, na forma do art. 32, pelo qual os sistemas inteligentes que aumentem a eficiência e a qualidade das atividades não poderão comprometer a autonomia dos usuários internos, permitindo-se uma revisão detalhada dos produtos gerados e dos dados processados, como consectário da transparência.

Ainda sobre esta preocupação, a resolução expressa por seu art. 34 que os sistemas computacionais utilizados no Poder Judiciário, obrigatoriamente, exigirão a supervisão humana e permitirão a modificação pelo juízo natural. Essa providência, acertadamente,

obsta que os sistemas automatizados sejam utilizados para a tomada de decisões exclusivamente, assegurando-se que a IA sirva como instrumento de apoio, e não como um equivalente jurisdicional.

Como bem salientam Rodrigues, Mendonça e Nóvoa (2025), “ao garantir a supervisão humana, evita-se que decisões tomadas exclusivamente por algoritmos resultem em vieses ou falhas que comprometam a equidade do processo judicial, preservando o caráter humano, imparcial e contextual do julgamento”.

As garantias de supervisão humana pretendem evitar vieses e prevenir a discriminação, havendo uma determinação de avaliação e validação contínua das soluções, nos termos do art. 8º:

Deverão ser implementadas medidas preventivas para evitar o surgimento de vieses discriminatórios, incluindo a validação contínua das soluções de IA e a auditoria ou monitoramento de suas decisões ao longo de todo o ciclo de vida da aplicação, para garantir que as soluções de IA continuem em conformidade com os princípios da igualdade, pluralidade e não discriminação, com relatórios periódicos que avaliem o impacto das soluções no julgamento justo, imparcial e eficiente.

Com a identificação de um viés, a resolução estabelece a pronta aplicação de medidas corretivas, “que podem incluir a suspensão temporária ou definitiva da solução, com o objetivo de garantir seu alinhamento com os princípios éticos e legais, promovendo um ambiente de justiça e imparcialidade” (Rodrigues; Mendonça; Nóvoa, 2025).

O uso responsável de IAs generativas também foi objeto de atenção da nova regulamentação. Os modelos de linguagem de larga e pequena escala, como o ChatGPT e Deepseek, embora possam ser utilizados no Judiciário, deverão, preferencialmente, ter acesso monitorado e habilitado pelos tribunais (art. 19, IV):

É vedado o uso de LLMs e sistemas de IA generativa de natureza privada ou externos ao Judiciário para processar, analisar, gerar conteúdo ou servir de suporte a decisões a partir de documentos ou dados sigilosos ou protegidos por segredo

de justiça, nos termos da legislação aplicável, salvo quando devidamente anonimizados na origem ou quando forem adotados mecanismos técnicos e procedimentais que garantam a efetiva proteção e segurança desses dados e de seus titulares

Complementarmente, proibiu-se o uso desses sistemas para processar, analisar, gerar ou auxiliar decisões com base em documentos sigilosos ou protegidos por segredo de justiça, excepcionando-se a regra tão somente quando houver a anonimização dos dados, na forma da LGPD e nas diretrizes da ANPD, ou com a adoção de métodos que permitam a proteção adequada dos dados.

O arranjo protecionista desenhado pela nova regulamentação do CNJ dá continuidade, com louvor, à missão de tutelar os direitos dos jurisdicionados, privilegiando a tutela daqueles que são fundamentais, como os dados pessoais, o devido processo, a igualdade, o acesso à justiça e a transparência. Esta, seguramente, como suporte primevo da explicabilidade, “imprescindível para a consolidação de um sistema acessível, com equidade na contestação e apresentação de alternativas” (Rodrigues; Mendonça; Nóvoa, 2025).

## 2 A INTELIGÊNCIA ARTIFICIAL EXPLICÁVEL

### 2.1 Definição e Importância da Inteligência Artificial ExPLICÁVEL

A transparência é um dos princípios essenciais da Administração Pública (art. 37, *caput*, Constituição Federal). Sua aplicação no Poder Judiciário assumiu uma posição central nas arquiteturas decisórias, que exigem a demonstração da *ratio decidendi*, na linha desenhada pelo ordenamento. A transparência aplicada à utilização da IA busca concretizar um “acesso claro às informações e processos de tomada de decisão, explicando como as decisões são feitas e quais critérios e algoritmos (sequências de instruções lógicas) são aplicados” (Pádua; Lorenzetto, 2024).

Em conferência realizada pelo *Future of Life Institute* (2017), desenvolveram-se os princípios de Asilomar, consistentes em

diretrizes que visam garantir que o desenvolvimento e a aplicação da Inteligência Artificial (IA) sejam realizados de forma benéfica e segura para a sociedade. Eles orientam que a utilização de sistemas em decisões judiciais (e da Administração Pública) deve ser acompanhada de explicações satisfatórias, suscetíveis à auditorias de agentes humanos, bem ainda que, em nome da transparência, acaso um sistema de IA cause danos, as soluções devem permitir o apontamento das razões para tanto (*Future of Life Institute*, 2017).

A inteligência artificial explicável (XAI, do inglês *Explainable Artificial Intelligence*) é um ramo da IA focado em desenvolver modelos cujas decisões possam ser facilmente compreendidas e interpretadas por humanos. No contexto do Poder Judiciário, a importância da XAI é particularmente relevante, pois ela promove a transparência e a responsabilidade (*accountability*) nas decisões automatizadas.

Sérgio Rodrigode Pádua(2021)qualifica o direito à explicabilidade como elemento da transparência das decisões tomadas por IA, cujo conteúdo pode ser extraído da Recomendação 449 da OCDE, na *First Version of a Draft Text of a Recommendation on the Ethics of Artificial Intelligence* da UNESCO, dos *Ethical Guidelines for Trustworthy AI* da União Europeia, do art. 20 da Lei Geral de Proteção de Dados (Lei 13.709/2018), da Resolução 332/2020 do Conselho Nacional de Justiça e, agora, da Resolução 615/2025.

O CNJ, como visto, preocupa-se com a transparência na utilização de ferramentas e sistemas de IA no Judiciário. A nova regulamentação dedicou o inciso XVIII do artigo 4º para definir a explicabilidade como sendo a “compreensão clara, sempre que tecnicamente possível, de como as “decisões” são tomadas pela IA”.

Como foi erigida à condição de princípio, os aspectos de governança da resolução também são vetorizados pela explicabilidade. É o caso das soluções classificadas em alto risco que, antes de sua produção, devem adotar medidas que viabilizem explicabilidade adequada e de medidas para disponibilizar informações em linguagem simples e acessível (art. 13, inciso VII).

No ambiente jurídico, onde as decisões podem afetar significativamente a vida das pessoas, compreender como essas decisões são tomadas é crucial. A falta de explicabilidade pode levar à desconfiança no sistema judicial e a questionamentos sobre a equidade e justiça das decisões. A IA explicável assegura que as partes interessadas — juízes, advogados e os próprios litigantes — possam entender a lógica e o raciocínio por trás das conclusões geradas por sistemas de IA.

Os esforços, portanto, são para explicar como um modelo de IA chegou ao resultado (*output*) verificado ou para auxiliar na interpretação deste, passando por processo de *accountability*. Nesse sentido, Pádua (2021):

[...] a explicabilidade no desenvolvimento de modelos de IA judicial inclui a transparência no processo construção do modelo algorítmico, o que consiste na definição do objetivo (primeira fase), na escrita do código algorítmico (coding – segunda fase) e na implementação do modelo (terceira fase). Nessa linha, a explicabilidade do processo de construção de modelos de IA tem pressupostos característicos que são distintos da explicabilidade das saídas (*outputs*) e resultados (*outcomes*) propiciados por um modelo de IA.

A Organização das Nações Unidas também destacou a importância da explicabilidade ao aprovar a sua primeira resolução sobre o assunto, em 2024:

4. Convoca os Estados Membros e convida outros interessados a tomar medidas para cooperar e fornecer assistência aos países em desenvolvimento para acesso inclusivo e equitativo aos benefícios da transformação digital e sistemas de inteligência artificial seguros, protegidos e confiáveis, incluindo por meio de: (...) (k) Promovendo transparência, previsibilidade, confiabilidade e compreensibilidade (grifo nosso) ao longo do ciclo de vida de sistemas de inteligência artificial que tomam ou apoiam decisões que impactam os usuários finais, incluindo o fornecimento de notificações e explicações (grifo nosso), e promovendo supervisão humana, como, por exemplo, através da revisão de decisões automatizadas e processos relacionados ou, quando apropriado e relevante, alternativas de decisão humana ou reparação eficaz e responsabilidade para aqueles adversamente impactados por decisões automatizadas de sistemas de inteligência artificial.

A explicabilidade é vital para mitigar riscos de vieses algorítmicos e de discriminações que possam perpetuar ou exacerbar desigualdades sociais. Quando modelos de IA são opacos, eles podem tomar decisões baseadas em correlações espúrias ou dados enviesados, resultando em resultados injustos, de modo que a XAI atua como uma salvaguarda, permitindo que decisões sejam auditadas e questionadas de forma eficaz, garantindo que todas as partes possam exercer seus direitos de forma plena e assegurando a legitimidade jurídica dos sistemas de IA utilizados pelo Judiciário. Quanto à opacidade, Pecego e Teixeira (2024) advertem:

A opacidade algorítmica é caracterizada pela impossibilidade de se obter uma explicação completa sobre o funcionamento interno, recomendações, decisões e previsões de sistemas de IA descritos como “black box”. As causas dessa opacidade incluem: i) proteção à propriedade intelectual e ao segredo de negócio em sistemas proprietários, onde as empresas desenvolvedoras mantêm os algoritmos em segredo para proteger seu investimento; ii) a necessidade de conhecimentos técnicos específicos para compreender sistemas complexos de machine learning, que envolvem matemática avançada e ciência da computação; e iii) a opacidade intrínseca a modelos computacionais complexos, como redes neurais profundas, que sofrem ajustes contínuos durante o aprendizado e operação. Embora essa complexidade permita maior precisão ao identificar padrões complexos, acaba por tornar os modelos mais difíceis de interpretar. Por outro lado, modelos mais simples e interpretáveis, como regressões lineares, oferecem transparência, mas podem ser menos precisos ao lidar com dados não estruturados, resultando em desempenho inferior, se comparados aos modelos mais avançados.

Ao concluir densa investigação sobre as múltiplas manifestações do direito à explicabilidade<sup>7</sup>, Pádua e Lorenzetto (2024) apresentam os seguintes eixos de afirmação do princípio:

- a) a oportunidade de conhecimento para o indivíduo cujos dados pessoais foram tratados no processo de modelagem da IA, para que ele possa entender como seus dados são utilizados;
- b) o fornecimento de explicações comprehensíveis sobre a lógica, etapas e processos do desenvolvimento de sistemas de IA que

<sup>7</sup> Os autores refletem que o direito à explicabilidade apresenta duas dimensões. Uma, de interpretabilidade, que diz respeito aos processos computacionais da IA e aos resultados gerados, facilitando a compreensão por parte dos profissionais da tecnologia; e a explicabilidade em sentido estrito, que se refere ao desenvolvimento de modelos em linguagem natural ou outras formas de representação que sejam facilmente compreensíveis pelos destinatários humanos das decisões auxiliadas por IA.

auxiliam a tomada de decisão por autoridades públicas (da Administração Pública e do Judiciário), sempre respeitando os limites técnicos e o máximo alcance viável conforme a tecnologia utilizada;

c) a possibilidade de profissionais técnicos (com conhecimento em Ciência da Computação) desafiarem o sistema de IA por meio de informações que permitam a interpretabilidade do sistema computacional.

Para Nunes e Andrade (2023), no contexto da IA, explicabilidade significa compreender melhor os motivos e detalhes por trás de uma decisão algorítmica, razão pela qual, enquanto a opacidade cria uma “caixa preta”, limitando a compreensão humana de decisões de um sistema de IA, a explicabilidade produziria uma “caixa de vidro”, permitindo a visualização dos processos por trás da lógica algorítmica.

Dessas bases teóricas e pelo papel que desempenha, diante do aumento e da necessária aplicação de técnicas de IA para auxiliar na tomada de decisões no âmbito do Judiciário, a explicabilidade deve ser entendida como um princípio e como um direito fundamental, necessário ao enfrentamento da opacidade e de vieses algorítmicos.

## 2.2 Ferramentas e Técnicas para a Explicabilidade em Sistemas de IA

Para alcançar a explicabilidade em sistemas de IA, várias ferramentas e técnicas foram desenvolvidas, divididas principalmente em dois enfoques: o design de modelos de caixa branca e as técnicas de interpretação pós-modelagem.

### 2.2.1 Modelos de Caixa Branca

Os modelos de caixa branca são projetados desde o início para serem interpretáveis. Um exemplo típico são as árvores de decisão, que oferecem um caminho visual claro e comprehensível das decisões. Esses modelos, embora possam ser menos precisos em comparação com modelos de caixa preta complexos (como redes

neurais profundas), são frequentemente preferidos em contextos onde a transparência é crítica.

Outro exemplo são os modelos lineares generalizados, que, embora simples, fornecem coeficientes claros que indicam a contribuição de cada variável na decisão final. Isso oferece uma maneira direta de entender o impacto dos diferentes fatores considerados pelo modelo.

### 2.2.2 Técnicas de Interpretação Pós-Modelagem

Para modelos complexos, onde a interpretação direta não é possível, existem técnicas de interpretação pós-modelagem. Essas técnicas são aplicadas após o desenvolvimento dos modelos para gerar explicações compreensíveis. Um exemplo disso é o método LIME (*Local Interpretable Model-agnostic Explanations*), que cria modelos interpretáveis locais para explicar previsões individuais de modelos complexos. Ele age como uma lente de aumento sobre partes específicas de dados, ajudando a entender como um modelo chegou a uma decisão particular.

Alguns modelos de aprendizado de máquina apresentam elevada complexidade, exigindo o uso de métodos específicos para que suas decisões possam ser compreendidas. É nesse contexto que surgem as técnicas de explicabilidade *post-hoc* (pós-modelagem), voltadas à interpretação de sistemas já treinados, especialmente aqueles cuja estrutura interna não é naturalmente transparente. Essas técnicas podem ser classificadas como gerais (ou agnósticas) — aplicáveis a diversos modelos — ou específicas, desenvolvidas para interpretar algoritmos individuais.

Dentre as técnicas gerais, que interessam ao presente estudo, destacam-se as explicações textuais, que descrevem o funcionamento do modelo em linguagem acessível, sem recorrer ao código-fonte. As explicações por simplificação, por sua vez, extraem regras representativas para tornar o processo decisório mais inteligível. Já as explicações visuais recorrem a gráficos ou esquemas — como os

gráficos ICE — para ilustrar as relações entre variáveis e previsões (Nunes; Andrade, 2023).

Outro método importante é o de atribuição de relevância, que quantifica a influência de cada variável na decisão algorítmica, sendo útil, inclusive, na detecção de possíveis vieses. As explicações pontuais concentram-se em trechos específicos do modelo, oferecendo uma visão localizada porém eficaz.

Ademais, há o método o SHAP (*SHapley Additive exPlanations*), que fornece valores de Shapley para cada característica, indicando sua influência na previsão de um modelo. Baseando-se na teoria dos jogos, o SHAP oferece uma maneira consistente e teórica de atribuir importância às variáveis de entrada, tornando-o uma ferramenta poderosa para a compreensão de modelos complexos.

Por fim, tem-se a técnica de exemplificação, que opera por meio da apresentação de casos similares aos analisados, facilitando a compreensão por meio da analogia. No magistério de Dierle Nunes e Otávio Morato (2023), nesse método “o algoritmo extrai amostras de dados análogos ou similares aos resultados gerados pelo modelo, permitindo que o usuário compreenda, por analogia, as relações e correlações encontradas pelo modelo algorítmico”.

A resolução do CNJ, regulamentadora do uso da IA no Judiciário, não define, expressamente, uma metodologia para a explicabilidade a ser adotada nacionalmente. Ela se limita a indicar que os modelos de IA deverão “possuir mecanismos de explicabilidade, sempre que tecnicamente possível, de modo que suas decisões e operações sejam compreensíveis e auditáveis pelos operadores judiciais” (art. 22, §3º). Sem prejuízo, deve-se somar isso ao terreno da auditabilidade, cujo foco parece residir no resultado das decisões geradas por IA.

Por consequência, infere-se que “se opta por uma abordagem preferencialmente decomposicional”, uma vez que se concentra em

explicar ou replicar “raciocínio interno do modelo, permitindo que o magistrado supervisione efetivamente o processo decisório” (Pecego; Teixeira, 2024).

Nessa senda,

A abordagem decomposicional tenta explicar ou replicar o raciocínio interno do modelo. Uma forma direta dessa abordagem é a revelação do código-fonte do modelo. Ocorre que o conhecimento do código fonte muitas vezes não é suficiente para se obter uma explicabilidade suficiente pois, conforme visto no tópico relativo às possíveis causas da opacidade, mesmo com a divulgação dos detalhes de um algoritmo, sua compreensão pode ser difícil para quem não possui o conhecimento técnico necessário. Além disso, mesmo programadores não conseguem extrair o motivo pelo qual algoritmos de machine learning fazem determinadas recomendações (Pecego e Teixeira, 2024).

Outra forma de abordagem decomposicional é a criação de um modelo substituto, “que analisa pares de entrada e saída, sem necessidade de acesso aos pesos internos do modelo original”. Exemplificativamente, “uma árvore de decisão pode ser usada para refletir os cálculos de um modelo *black box*, permitindo compreender quais fatores influenciam suas decisões” (Pecego; Teixeira, 2024). Isso, todavia, não obsta a adoção de outras técnicas possíveis, tendentes à assegurar os critérios de governança pretendidos pelo CNJ.

Em suma, a implementação de IA explicável no sistema judiciário não apenas melhora a transparência, mas também sustenta a confiança pública no uso da tecnologia. À medida que o Poder Judiciário brasileiro, incluindo tribunais como o TJRO, avança em suas iniciativas de digitalização e automação, adotar e aprimorar técnicas de XAI será crucial para equilibrar inovação tecnológica com justiça e equidade. Isso garantirá um judiciário mais eficiente e confiável, capaz de lidar com os desafios da era digital de maneira responsável.

### 3 ESTUDO DE CASO: TRIBUNAL DE JUSTIÇA DE RONDÔNIA

#### 3.1 Iniciativas e Implementações

O Tribunal de Justiça de Rondônia (TJRO) tem se destacado no cenário nacional pela adoção proativa e inovadora de tecnologias de inteligência artificial (IA). Essa postura vanguardista é evidenciada por várias iniciativas que visam integrar soluções de IA para aprimorar a eficiência e a acessibilidade dos serviços judiciais.

Entre as principais iniciativas está o desenvolvimento de sistemas de automação para a triagem de processos, que utilizam algoritmos de IA para categorizar e priorizar casos, permitindo que sejam encaminhados de forma mais eficiente para as varas competentes.

Outro projeto notável é a utilização de ferramentas de análise preditiva, que têm potencial para auxiliar na identificação de padrões em processos judiciais, ajudando a prever desfechos com base em decisões anteriores e características dos casos analisados. Essas ferramentas não apenas aceleram a resolução de casos repetitivos, mas também ajudam na alocação de recursos judiciais, otimizando a carga de trabalho dos juízes e servidores.

#### 3.1 Impactos e Resultados

As implementações de IA no TJRO têm gerado impactos significativos tanto internamente quanto na percepção pública do tribunal. Internamente, as tecnologias introduzidas resultaram em maior eficiência processual, com a redução do tempo de tramitação dos processos e melhoria na gestão do acervo processual. O uso de IA para automatizar tarefas repetitivas liberou recursos humanos para se concentrarem em atividades de maior complexidade e valor jurídico.

Na percepção pública, essas inovações têm contribuído para um aumento na confiança e na satisfação dos usuários do sistema judicial. A transparência aprimorada e a maior acessibilidade das informações processuais fortalecem a imagem do TJRO como uma instituição comprometida com a inovação e a modernização. Além disso, a percepção de um judiciário mais ágil e acessível tem potencial para aumentar a participação da sociedade no processo judicial, fortalecendo o papel do tribunal como um facilitador de justiça.

### **3.2 Desafios Locais**

Apesar dos avanços, a implementação de IA no TJRO enfrenta desafios específicos, muitos dos quais estão ligados à necessidade de garantir a explicabilidade e a ética no uso dessas tecnologias. Um dos principais desafios é desenvolver sistemas de IA que sejam transparentes e compreensíveis para todos os usuários, evitando a “caixa preta” das decisões automatizadas. Isso requer um investimento contínuo em soluções de inteligência artificial explicável (XAI) que possam assegurar que as decisões automatizadas sejam auditáveis e justificáveis.

Além disso, há desafios relacionados à capacitação e ao treinamento de servidores e magistrados para operarem e interpretarem corretamente as ferramentas de IA. A integração de novas tecnologias demanda um esforço considerável em termos de treinamento e mudança cultural dentro da instituição, assegurando que todos os envolvidos estejam preparados para utilizar as novas ferramentas de forma eficaz e responsável.

Outro desafio significativo é a proteção de dados sensíveis. A conformidade com a Lei Geral de Proteção de Dados (LGPD) é crucial para garantir que a implementação de IA respeite a privacidade dos indivíduos e proteja as informações processadas contra acessos não autorizados.

Para enfrentar esses desafios, o TJRO tem investido em pesquisa e em parcerias com empresas tecnológicas, buscando desenvolver soluções personalizadas que atendam às necessidades específicas do tribunal. Além disso, a instituição promove *workshops* e seminários para educar e engajar seus colaboradores nos novos paradigmas tecnológicos.

Em conclusão, o TJRO representa um estudo de caso exemplar na adoção de IA no sistema judiciário brasileiro, demonstrando como a inovação tecnológica pode ser aplicada de forma responsável para melhorar a prestação de serviços judiciais. Ao mesmo tempo, continua a enfrentar e superar desafios típicos dessa transformação digital, servindo como modelo para outras jurisdições no país.

#### 4. CONSIDERAÇÕES FINAIS

Este artigo analisou a crescente incorporação da inteligência artificial (IA) no Poder Judiciário brasileiro, com um foco particular no conceito de IA explicável e suas práticas no Tribunal de Justiça de Rondônia (TJRO). Ao longo dos três capítulos, procurou-se descrever o impacto transformador que essas tecnologias exercem sobre o sistema judicial, ao mesmo tempo em que se reconhecem os desafios éticos e operacionais que acompanham sua adoção.

No primeiro capítulo, foi discutido o papel da IA no Judiciário, destacando como sua evolução trouxe benefícios como eficiência e redução de custos, mas também introduziu desafios, especialmente no que tange às questões de transparência e justiça. A importância de um judiciário ágil e eficiente não pode ser subestimada, mas deve ser equilibrada com a necessidade de manter a justiça como um princípio fundamental.

O segundo capítulo explorou a inteligência artificial explicável (XAI), enfatizando sua relevância no contexto judicial, onde a clareza e a justificabilidade das decisões são de suma importância. As ferramentas e técnicas apresentadas, como modelos de caixa branca e métodos de interpretação pós-modelagem, são essenciais para assegurar que as decisões de IA sejam compreensíveis e auditáveis, promovendo a *accountability* e mantendo a confiança pública.

No estudo de caso apresentado no terceiro capítulo, o TJRO ilustrou como a adoção inovadora de IA pode beneficiar tanto a eficiência interna quanto a percepção pública do sistema judicial. As iniciativas do TJRO oferecem uma visão prática de como a IA pode ser implementada de forma a melhorar a operacionalidade do tribunal, ao mesmo tempo em que enfrentam desafios relativos à explicabilidade, treinamento e proteção de dados.

Em suma, a incorporação da IA no TJRO não só demonstra o potencial dessas tecnologias para reformular os processos judiciais, mas também serve como um lembrete da necessidade de implementar práticas responsáveis e éticas. A busca pela inovação deve ser orientada por princípios que assegurem justiça, equidade e transparência. O TJRO, ao enfrentar desafios locais e desenvolver soluções adaptativas, oferece um modelo valioso que pode inspirar outros tribunais a seguir um caminho similar.

O futuro do Judiciário brasileiro, apoiado por tecnologias de IA, depende da capacidade das instituições de equilibrar inovação com responsabilidade. Para tanto, é essencial continuar investindo em pesquisa, desenvolvimento e treinamento, de modo a garantir que a tecnologia sirva como uma aliada no fortalecimento do sistema judicial. O caminho trilhado pelo TJRO é promissor, mas o esforço deve ser contínuo para consolidar os ganhos em eficiência e justiça, assegurando que o Judiciário brasileiro se mantenha à frente das transformações tecnológicas globais.

## REFERÊNCIAS

BRASIL. Conselho Nacional de Justiça. Resolução n. 332, de 18 de dezembro 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 Eletrônico, Brasília, DF, 21 dez. 2020.

BRASIL. Conselho Nacional de Justiça. Resolução n. 615, de 11 de março de 2025. Estabelece diretrizes para o desenvolvimento, utilização e governança de soluções desenvolvidas com recursos de inteligência artificial no Poder Judiciário. Diário da Justiça Eletrônico, Brasília, DF, 31 mar. 2025.

CERRI, Ricardo; CARVALHO, André Carlos Ponce de Leon Ferreira de. Aprendizado de máquina: breve introduções e aplicações, Cadernos de Ciência e Tecnologia, 2017.

COSTA E.; SIMÕES A., Inteligência artificial: fundamentos e aplicações, Editora FCA, 2008.

DOMINGOS, Pedro. A revolução do algoritmo mestre: como a aprendizagem automática está a mudar o mundo, Manuscrito, 2017.

FACELI, Katti et al. Inteligência artificial: uma abordagem de aprendizado de máquina, LTC, 2021.

FUTURE OF LIFE INSTITUTE. *Asilomar AI Principles*. 2017. Disponível em: <https://futureoflife.org/ai-principles>. Acesso em: 10 abr. 2025.

HAWKING, Stephen; RUSSELL, Stuart; TEGMARK, Max; WILCZEK, Frank. Transcendence looks at the implications of artificial intelligence, but are we taking AI seriously enough? Independent, 2014. Disponível em: <https://www.independent.co.uk/news/science/stephen-hawking-transcendence-looks-at-the-implications-of-artificial-intelligence-but-are-we-taking-ai-seriously-enough-9313474.html>. Acesso em: 18 abr. 2025.

LORENZETTO, Bruno Meneses; PÁDUA, Sérgio Rodrigo de. O direito fundamental à explicabilidade da inteligência artificial utilizada em decisões estatais. Revista da AGU, v. 23, n. 66, p. 25-45, 2024.

NUNES, Dierle José Coelho; ANDRADE, Otávio Morato de. O uso da inteligência artificial explicável enquanto ferramenta para compreender decisões automatizadas: possível caminho para aumentar a legitimidade e confiabilidade dos modelos algorítmicos? Revista Eletrônica do Curso de Direito da UFSM [Recurso Eletrônico], Santa Maria, RS, v.18, n.1, 2023.

NUNES PECEGO, D.; JANNY TEIXEIRA, Raphael Lobato Collet, R. (2024). Inteligência artificial no Judiciário: da opacidade à explicabilidade das decisões judiciais. Revista Da Faculdade De Direito Da UERJ - RFD, (43), 1-22. <https://doi.org/10.12957/rfd.2024.87850>.

RODRIGUES, Carla; MENDONÇA, Eduardo; NÓVOA, Natasha. IA e as mudanças no judiciário brasileiro. Data Privacy Brasil, 2025. Disponível em: <https://www.dataprivacybr.org/wp-content/uploads/2025/03/IA-e-as-mudancas-no-judiciario-brasileiro-Data-Privacy-Brasil.pdf>. Acesso em: 18 abr. 2025.

SILVA, Ivan Nunes da. Redes neurais artificiais para engenharia e ciências aplicadas: fundamentos teóricos e aspectos práticos, Artliber, 2016.

PÁDUA, Sérgio Rodrigo de. Explicabilidade e inteligência artificial judicial. Anais do EVINCI – UniBrasil, Curitiba, v.7, n.1, p. 338-338, out. 2021.