

Revista Brasileira de Ciências Sociais Aplicadas

ISSN 3085-8151

vol. 2, n. 1, 2026

... ARTIGO 1

Data de Aceite: 02/01/2026

INTELIGÊNCIA ARTIFICIAL NO JUDICIÁRIO BRASILEIRO: RESPONSABILIDADES, TRANSPARÊNCIA E GARANTIAS PROCESSUAIS EM UMA REVISÃO DE ESCOPO (2019–2026)

Emiliano Peggion de Carvalho Navarro

Possui graduação em Ciências Sociais (Unesp) e Direito (Unirv), especialização em mediação e conciliação e direito constitucional, mestre e doutorando em Estudos de Cultura Contemporânea pela UFMT. Coordenador do Núcleo de Prática Jurídicas da Faculdade Anhanguera e Professor na Faculdade Anhanguera de Tangará da Serra, Unemat de Barra do Bugres e efetivo do Estado de Mato Grosso.

<http://lattes.cnpq.br/4650569202876327>

Edilene Maria Torquato Villar

Possui graduação em Direito (UNOESTE), Especialização em Direito Processual Civil (ESMP). Professora e Coordenadora do Curso de Direito na Faculdade Anhanguera de Tangará da Serra/Mato Grosso.

<http://lattes.cnpq.br/3661522854716464>

Eliziane Fernanda Navarro Peggion

Doutora em Letras e Linguística (Universidade Federal de Goiás), Mestre em Estudos Literários (Universidade do Estado do Mato Grosso). Graduada em Direito (Universidade do Estado do Mato Grosso), Letras, Português, Espanhol e respectivas Literaturas (Universidade do Estado do Mato Grosso) e Letras – Inglês (Instituto Prominas). Atua como professora na Faculdade Anhanguera e como Professora Formadora de Gestão na SEDUC/MT

<http://lattes.cnpq.br/6725500562364847>



Todo o conteúdo desta revista está licenciado sob a Licença Creative Commons Atribuição 4.0 Internacional (CC BY 4.0).

Resumo: A incorporação de sistemas computacionais de apoio à decisão e de automação de tarefas no sistema de justiça tem se intensificado no Brasil, reabrindo debates sobre controle, legitimidade e proteção de direitos em um cenário de digitalização acelerada. Este artigo tem como objetivo mapear e organizar, por meio de revisão de escopo, a literatura acadêmica e técnico-jurídica publicada entre 2019 e 2026 acerca de três eixos críticos: (i) responsabilidades (institucionais, funcionais e eventualmente civis) relacionadas ao desenvolvimento, contratação e uso de soluções algorítmicas; (ii) transparência e auditabilidade (incluindo explicabilidade, registro de decisões e possibilidade de escrutínio público e processual); e (iii) garantias processuais, com ênfase em contraditório, ampla defesa, motivação e controle recursal em ambientes influenciados por tecnologia. A metodologia compreende estratégia de busca estruturada em bases acadêmicas e repositórios, seleção por critérios de inclusão e exclusão previamente definidos, e extração padronizada de dados para categorização temática e síntese narrativa organizada em mapas e quadros analíticos. Como resultado, a produção é sistematizada em tipologias de uso (triagem, classificação, recomendação, gestão de acervos e elaboração de minutas), riscos recorrentes (opacidade, vieses, dependência institucional e assimetrias informacionais) e propostas de salvaguardas (governança, documentação técnica, deveres de informação, protocolos de revisão humana e mecanismos de responsabilização). Conclui-se com uma agenda de pesquisa e recomendações para compatibilizar inovação tecnológica com garantias fundamentais e confiança institucional.

Palavras-chave: Automação decisória; Devido processo legal; Transparência algorítmica; Accountability; Governança digital.

Construção histórica das IAs: do simbolismo ao aprendizado de máquina

A expressão “inteligência artificial” consolidou-se no século XX, com marco recorrente na reunião de Dartmouth (1956), frequentemente apontada como momento fundador do campo, quando o termo passa a designar um programa científico de investigar e construir máquinas capazes de desempenhos associados à cognição humana. (UFOPA, s.d.) Nesse período e nas décadas seguintes, grande parte do esforço se organizou em torno da chamada IA simbólica, baseada em lógica, regras e manipulação de símbolos, com expectativa de que o raciocínio pudesse ser formalizado e executado por programas. (UFOPA, s.d.)

A virada contemporânea, porém, é marcada pela centralidade do aprendizado de máquina: em vez de depender primariamente de regras explicitadas por especialistas, modelos passam a “aprender” padrões a partir de dados, com desempenho ajustado por treino e validação. A literatura de divulgação histórica registra, por exemplo, Arthur Samuel como pioneiro da ideia de sistemas que melhoram desempenho por experiência, ao desenvolver um programa que aprimorava seu jogo de damas ao longo do tempo. (IBM, 2024) Esse deslocamento altera o tipo de controle possível: se, em muitos sistemas simbólicos, a justificativa pode ser rastreada a regras legíveis, nos modelos estatísticos complexos a justificativa tende a se dispersar em parâmetros e correlações, dificultando a reconstrução do “porquê” de uma saída.

Daí emerge um tema que atravessa o debate jurídico-institucional atual: a exigência de interpretabilidade/explicabilidade como condição de legitimidade em usos sensíveis. Doshi-Velez e Kim observam que, com a ubiquidade do aprendizado de máquina, cresceu o interesse por modelos interpretáveis, mas também permanece pouca convergência sobre o que “interpretabilidade” significa e como medi-la. (Doshi-Velez; Kim, 2017) Em documentos técnicos de referência, o NIST sintetiza que sistemas explicáveis devem “deliver accompanying evidence or reasons” (entregar evidências ou razões) para seus resultados e processos, entre outros princípios, reforçando que explicabilidade não é atributo meramente estético, mas propriedade funcional ligada a confiança e controle. (Phillips et al., 2021)

Esse pano de fundo histórico importa para o Judiciário por um motivo simples: a dificuldade (ou possibilidade) de explicar resultados não é um detalhe técnico; ela molda a forma como responsabilidades são atribuídas, como o controle institucional ocorre e como garantias processuais podem ser exercidas. Quando a “racionalidade” do sistema não é acessível por regras claras, o debate jurídico precisa migrar do ideal de transparência total (como se todo raciocínio fosse plenamente legível) para arquiteturas de governança, auditoria, registro e contestação, capazes de produzir controle prático mesmo em cenários de complexidade técnica. (Phillips et al., 2021)

A entrada da IA no Judiciário brasileiro: infraestrutura, gestão e padronização

No contexto brasileiro, a adoção de IA no Judiciário deve ser entendida como parte de uma política de transformação digital que combina racionalização de fluxos, padronização de rotinas e busca por ganhos de escala em tarefas repetitivas. Um marco concreto desse movimento é a Plataforma Sinapses, articulada pelo CNJ no âmbito do Programa Justiça 4.0 e concebida como catálogo e ambiente de treinamento/consumo de modelos para órgãos do sistema de justiça. (PNUD, 2023) Em divulgação institucional, registra-se que há “Cento e cinquenta modelos de Inteligência Artificial (IA) ativos” depositados na plataforma, produzidos por dezenas de tribunais e conselhos, indicando adoção capilar e crescente. (PNUD, 2023)

O próprio desenho do Sinapses sugere uma mudança qualitativa: não se trata apenas de cada tribunal desenvolver sua solução de forma isolada, mas de organizar reuso, compartilhamento e padronização mínima, com impactos sobre governança e rastreabilidade. (PNUD, 2023) Além disso, a experiência institucional reporta usos típicos voltados à identificação de classe/assunto no peticionamento e à indicação de precedentes/temas relevantes, isto é, aplicações que influenciam o “caminho” do processo, mesmo quando não substituem o julgamento do mérito. (PNUD, 2023)

Esse tipo de uso evidencia que a IA, no Judiciário, frequentemente opera como tecnologia de gestão: ela reorganiza prioridades, agrupa demandas, sinaliza repetitividade, sugere caminhos e reduz custos de classificação. (PNUD, 2023) O efeito ju-

rídico desse arranjo não é neutro, porque gestão do tempo e do foco institucional é também gestão de direitos em ambientes marcados por assimetria informacional e desigualdade de capacidades entre litigantes. Quando uma solução automatizada altera o fluxo (triagem, encaminhamento, temas, prevenção), ela altera, na prática, a forma como o acesso à ordem jurídica justa se materializa, o que exige debate normativo sobre publicidade, controle e contestabilidade.

Outro ponto sensível é a própria delimitação do que se chama “IA” no discurso institucional: muitas soluções operacionais podem ser automações determinísticas, enquanto outras são modelos probabilísticos treinados sobre bases locais. Essa distinção importa porque o risco jurídico não decorre apenas da automatização em si, mas do grau de indeterminação e do modo como o sistema generaliza padrões a partir de dados, podendo reproduzir distorções (“dirty data”) e vieses já inscritos em práticas institucionais. (AI Now Institute, s.d.) Por isso, o debate adequado precisa conectar tecnologia, dado e instituição: o que se automatiza, com qual dado, sob quais metas, e com quais formas de validação e controle contínuo.

Responsabilidades e transparência/auditabilidade: do “abrir a caixa-preta” ao desenho de controles

A discussão sobre responsabilidade em IA aplicada ao Judiciário não pode se restringir à pergunta “quem apertou o botão”. Em ambientes públicos, a cadeia de produção e operação é distribuída: há decisões de contratação, definição de requisitos, curadoria de dados, treino e retreino, parame-

trizações locais, integrações com sistemas processuais, uso cotidiano e governança de mudanças. (PNUD, 2023) Assim, quando ocorrem erros, vieses, perdas de rastreabilidade ou efeitos discriminatórios, a atribuição de responsabilidade tende a exigir uma análise de causalidade institucional: falha de projeto, falha de dado, falha de uso, falha de supervisão, ou combinação de todas.

É aqui que transparência e auditabilidade deixam de significar apenas “publicar código-fonte” e passam a significar capacidade de reconstrução do percurso decisório e de seus fundamentos. O Conselho da Europa, ao discutir transparência algorítmica e accountability, ressalta que processos ou decisões “taken or aided by algorithmic systems shall be identifiable and traceable” (tomados ou auxiliados por sistemas algorítmicos devem ser identificáveis e rastreáveis), além de exigir meios efetivos de contestação, mecanismos de supervisão e remédios. (Council of Europe, 2023) Essa formulação é particularmente útil para o Judiciário porque desloca o foco: a transparência relevante é aquela que habilita controle institucional e processual, inclusive por via recursal ou de reclamação, quando a tecnologia participa de encaminhamentos relevantes.

A literatura técnica também indica que explicabilidade é um conjunto de propriedades e não um “sim/não”. O NIST propõe princípios nos quais explicações devem ser acessíveis ao usuário, refletir corretamente o processo e vir acompanhadas de evidência/razões, sugerindo que o requisito de explicabilidade deve ser desenhado conforme finalidade e risco do uso. (Phillips et al., 2021) Em chave metodológica, Doshi-Velez e Kim defendem que a avaliação de interpretabilidade precisa ser rigorosa e situada, pois diferentes tarefas e usuários demandam dife-

rentes padrões de explicação. (Doshi-Velez; Kim, 2017) Para o Judiciário, isso implica reconhecer níveis: uma triagem administrativa pode admitir explicações padronizadas e testes de desempenho; já uma recomendação que afeta prioridade, distribuição ou escolha de precedente pode demandar rastreabilidade mais forte e registros mais detalhados.

Do ponto de vista de políticas públicas, a transparência útil também exige documentação: objetivo do modelo, limites de uso, fontes de dados, métricas de desempenho, taxas de erro, taxas de falso positivo/negativo, e logs que permitam auditoria retrospectiva. A experiência institucional brasileira com Sinapses é relevante porque, ao menos na apresentação pública, a plataforma é descrita como voltada ao “treinamento, à hospedagem e à auditabilidade de modelos de IA” para uso no Judiciário, sugerindo uma preocupação explícita com controle e padronização. (PNUD, 2023) Essa ideia de auditabilidade, porém, precisa ser concretizada em rotinas: quem audita, quando audita, com qual método, com quais consequências normativas (suspensão, retreino, correção, publicidade de achados).

Nesse ponto, a revisão de escopo proposta no artigo pode organizar o debate brasileiro em um mapa de “mecanismos de responsabilização” (accountability) e “mecanismos de transparência” (auditabilidade), distinguindo: (a) transparência informacional (divulgação ao público e às partes); (b) transparência processual (registro no caso concreto de uso e influência); (c) transparência técnica (documentação e testes); e (d) transparência institucional (governança, corregedorias, comitês, relatórios). O objetivo não é exigir um padrão único para toda aplicação, mas verificar como a literatura distribui exigências conforme risco e impac-

to, e quais lacunas aparecem entre proposta teórica e prática institucional.

Garantias processuais e contestabilidade: contraditório, motivação e controle em decisões mediadas por tecnologia

O núcleo mais sensível do debate aparece quando a IA influencia, direta ou indiretamente, decisões com efeitos sobre direitos. Mesmo quando um modelo não “decide o mérito”, ele pode impactar a trajetória do processo: prioridade, agrupamento, prevenção, sugestão de precedentes, padronização de minutas e triagens que condicionam o tempo e o tipo de atenção humana. (PNUD, 2023) Em tais situações, contraditório e ampla defesa não se esgotam na possibilidade formal de falar nos autos; exigem participação informada sobre critérios relevantes que afetaram o caso e meios reais de impugnar efeitos indevidos.

O debate internacional sobre risco de opacidade e devido processo é ilustrado por controvérsias envolvendo ferramentas algorítmicas em justiça criminal. No caso *Wisconsin v. Loomis*, discutiu-se a constitucionalidade do uso de um instrumento algorítmico de avaliação de risco na dosimetria/decisão, com alegações sobre ausência de explicação adequada e efeitos sobre o direito de defesa; a decisão aceitou o uso, mas a controvérsia alimentou forte crítica acadêmica sobre o alcance do devido processo quando o réu não consegue compreender o funcionamento interno do algoritmo. (Malgieri, 2018) Ainda que o contexto seja distinto do processo civil brasileiro, o problema estrutural é semelhante: como controlar uma ferramenta que influencia decisões se

os afetados não conseguem conhecer seus critérios de operação ou suas limitações.

Por isso, o direito à contestação — entendido como possibilidade de questionar decisões “tomadas ou auxiliadas” por sistemas — torna-se uma peça central de garantias em ambientes tecnológicos. O Conselho da Europa explicita que um quadro de transparência deve ser complementado por “effective means to contest” (meios efetivos de contestar) e por remédios de revisão judicial e não judicial. (Council of Europe, 2023) Em termos processuais, isso pode se traduzir em dever de informação sobre o uso da ferramenta, registro no processo (quando relevante), explicação mínima vinculada ao caso e preservação de via recursal apta a revisar o ato humano que se apoiou em recomendação algorítmica.

Outro risco é a deterioração da motivação. Se uma decisão ou encaminhamento passa a ser justificado apenas pelo argumento de autoridade técnica (“o sistema indicou”), a motivação pode se tornar aparente, isto é, incapaz de expor razões controláveis. Esse risco é mais intenso quando não há padrões de explicabilidade definidos para o tipo de tarefa, ou quando a organização premia produtividade sem exigir rastreabilidade. A literatura técnica, ao propor princípios de IA explicável, sugere que a explicação deve ser compreensível ao usuário e refletir corretamente o processo de geração do resultado, o que é compatível com a ideia jurídica de motivação controlável. (Phillips et al., 2021) Do ponto de vista probatório e argumentativo, a explicação não precisa revelar todo o modelo, mas precisa permitir que limitações e critérios sejam confrontados por sujeitos processuais e por instâncias revisoras.

Por fim, há o tema dos dados: se modelos são treinados com bases influenciadas

por assimetrias e vieses institucionais, o risco de reprodução automatizada dessas distorções aumenta. Um documento do AI Now Institute chama atenção para ferramentas que se apoiam em “dirty data”, isto é, dados derivados de versões enviesadas do mundo por práticas institucionais, o que pode comprometer justiça e confiabilidade. (AI Now Institute, s.d.) No Judiciário, isso exige que a literatura discuta não apenas “acurácia”, mas também critérios de qualidade do dado, curadoria, avaliação de impacto e governança de retreino, pontos que o próprio relato institucional do Sinapses indica como agenda de trabalho (curadoria de dados e avaliação de modelos). (PNUD, 2023)

Com esse conjunto, o desenvolvimento do artigo pode ser organizado, na revisão de escopo, em três camadas analíticas: (i) camada tecnológica (tipo de sistema, nível de explicabilidade, registros); (ii) camada institucional (governança, contratos, auditorias, responsabilização); e (iii) camada processual (dever de informação, contraditório, motivação e contestação). A ambição não é resolver, em abstrato, todas as controvérsias, mas mapear o estado do debate (2019–2026), localizar convergências (por exemplo, a necessidade de rastreabilidade e meios de contestação) e identificar lacunas que exigem pesquisa empírica e normatização mais precisa, especialmente em contextos de uso crescente e compartilhado de modelos no Judiciário brasileiro. (PNUD, 2023)

Se você concordar, eu adapto esse “desenvolvimento” para o formato de “Resultados e Discussão” da revisão de escopo, já incluindo: (a) um quadro de categorias de extração (responsabilidade, transparência, garantias); e (b) trechos com citações em ABNT autor-data prontos para sua lista de referências (com DOI/URL onde houver).

Eixo	Categoria (o que codificar)	Definição operacional	O que extrair (campos mínimos)	Observações de análise
Contexto/uso	Tipo de fonte	Natureza do material (acadêmico, técnico, normativo, institucional, jurisprudencial)	Autor/entidade; ano; tipo; local/país; link/DOI	Ajuda a separar “literatura científica” de documentos de implementação institucional.
Contexto/uso	Tipo de aplicação no Judiciário	Em que ponto do fluxo a IA/automação atua (ex.: triagem, classificação, recomendação, busca de jurisprudência, gestão de acervo, minutas)	Descrição do caso de uso; fase processual; usuário-alvo (servidor/magistrado/partes)	Exemplos concretos de aplicações no PJe e no Sinapses incluem identificação de classe/assunto e indicação de precedentes/temas.
Contexto/uso	Grau de impacto processual	Potencial de afetar direitos (baixo/médio/alto)	Justificativa do grau; ato afetado (priorização, distribuição, decisão, minuta etc.)	Pode ser articulado com abordagem “sensível ao risco” (risk-based).
Responsabilidade	Cadeia de responsabilização	Quem decide, desenvolve, contrata, parametriza e supervisiona	Órgão/atores; modelo de contratação; governança; logs	Útil para mapear “fragmentação” de responsabilidades (instituição–fornecedor–usuário).
Responsabilidade	Regime de responsabilidade discutido	Como a fonte enquadra responsabilização (administrativa, disciplinar, civil, contratual, estatal)	Tese central; fundamentos; limites; remédios	Em documentos institucionais, aparece com frequência a distinção entre apoio automatizado e responsabilidade final do magistrado.
Transparência/auditabilidade	Identificação do uso	Se a fonte defende dever de tornar identificável que houve “decisão/ato auxiliado por algoritmo”	Proposta concreta; quando informar; para quem informar	Parâmetro normativo útil: decisões “taken or aided” por sistemas algorítmicos devem ser “identifiable and traceable”. (CAPPELLO, 2023, p. 9).
Transparência/auditabilidade	Rastreabilidade (traceability)	Capacidade de reconstruir o percurso do sistema (entradas, versão do modelo, saídas, logs)	Logs; versionamento; auditoria; documentação	Facilita controle interno, externo e recursal.
Transparência/auditabilidade	Explicabilidade	Padrão de “explicação” exigido (ex.: justificativa textual; evidências; limites)	Tipo de explicação; destinatário; limitações	O NIST define como princípio que um sistema “deliver accompanying evidence or reasons” (entregue evidências/razões) para saídas/processos. (PHILLIPS et al., 2021, p. i). ^[59]

Garantias processuais	Contradi-tório/defesa informada	Se a parte pode conhecer e contestar o critério tecnicamente mediado	Mecanismos propostos; ônus de informação; momento	A contestabilidade é reforçada por padrões internacionais: “effective means to contest” e “effective remedies”. (CAPPELLO, 2023, p. 9).
Garantias processuais	Motivação/ controle	Se a motivação permanece controlável (evitar “o sistema disse”)	Critérios para fundamentar; papel do humano; revisão	Conecta explicabilidade a dever de motivação e revisibilidade. ^[59]
Garantias processuais	Risco de opacidade/ segredo	Como a fonte trata trade secrets/propriedade intelectual versus devido processo	Soluções propostas; níveis de acesso; auditoria	Debate análogo em justiça criminal algorítmica aparece em Loomis: questiona-se defesa plena sem acesso às “technicalities” do risco. (MALGIERI, 2018). ^[60]
Governança	Salvaguardas	Medidas sugeridas (comitês, auditorias, avaliação de impacto, retreino, curadoria de dados)	Quais; quem executa; periodicidade	No ecossistema Sinapses/ ColaboraJus há menção explícita a curadoria de dados, retreino e avaliação de modelos.

Quadro 1 – Matriz de extração e codificação (revisão de escopo, 2019–2026)

Resultados e Discussão

Tipologias de uso: IA como infraestrutura de gestão do fluxo

No debate recente, a IA no Judiciário aparece menos como “substituta do juiz” e mais como infraestrutura que reorganiza tarefas, prioridades e rotinas, incidindo na forma como a instituição vê, classifica e encaminha processos. Um exemplo expressivo é a Plataforma Sinapses, descrita como catálogo e espaço virtual de “treinamento e consumo” de modelos de IA pelos órgãos de Justiça. Em 2023, registrou-se a existência de “cento e cinquenta modelos” ativos depositados por 29 tribunais e conselhos, indicando uma política de escala e compartilhamento de soluções.

Esse “modo de entrada” importa porque desloca o problema jurídico do mérito decisório para a governança de fluxos: tria-

gem, classificação, recomendação, agrupamentos de casos e apoio à pesquisa. No PJe, há exemplos de implementação com modelos de identificação de classe e assunto no peticionamento inicial e de indicação de precedentes qualificados/temas, com a justificativa de simplificar fluxos e reduzir tarefas repetitivas. Ao mesmo tempo, o próprio discurso institucional estabelece uma linha de contenção: “Não houve geração de minuta de sentença”, e a fala enfatiza que “o magistrado... é o responsável pelas decisões e documentos produzidos”.

A literatura técnico-metodológica sobre interpretabilidade ajuda a compreender por que essa linha de contenção é instável: mesmo quando a máquina “só auxilia”, ela pode orientar decisões humanas e processos de compreensão. Doshi-Velez e Kim observam que sistemas de aprendizado de máquina “often guide processes of human understanding and decisions” (frequentemente

guiam processos de compreensão humana e decisões), o que torna a influência indireta juridicamente relevante. (DOSHI-VELEZ; KIM, 2017, p. 1).

Transparência, explicabilidade e auditabilidade: do “entender o modelo” ao “controlar o uso”

No conjunto de debates sobre governança algorítmica, transparência não é tratada como valor abstrato, mas como condição de possibilidade de controle e responsabilização. O IRIS Special do Observatório do Audiovisual Europeu (Conselho da Europa) formula isso de modo direto: “Transparency is a prerequisite for assessing the real impact of algorithmic systems... and for enforcing other rules aimed at establishing accountability”. (CAPPELLO, 2023). Em chave operacional, o mesmo documento afirma que processos ou decisões “taken or aided by algorithmic systems shall be identifiable and traceable...”, complementados por “effective means to contest... oversight mechanisms and effective remedies”. (CAPPELLO, 2023, p. 9).

Para o Judiciário, isso se desdobra em duas exigências práticas distintas: (i) identificabilidade/rastreabilidade do uso (saber que houve mediação tecnológica relevante, qual modelo/versão, com quais parâmetros e logs), e (ii) explicabilidade adequada ao destinatário (juiz, servidor, parte, instância revisora). No plano técnico, o NIST sintetiza quatro princípios para IA explicável, começando pelo requisito de que o sistema entregue “accompanying evidence or reasons” para saídas e/ou processos (princípio “Explanation”). (PHILLIPS et al., 2021, p.

i–ii). Também propõe que explicações sejam “Meaningful” (compreensíveis ao destinatário) e que tenham “Explanation Accuracy” (fidelidade ao processo do sistema), além de “Knowledge Limits” (o sistema operar apenas em condições para as quais foi desenhado e com confiança suficiente). (PHILLIPS et al., 2021, p. ii).

A utilidade desses princípios, no plano jurídico-processual, está em converter “explicabilidade” em requisitos verificáveis: evidência/razões, compreensão pelo usuário, fidelidade, e declaração de limites. Em sistemas de justiça, tais requisitos funcionam como antídotos contra justificações aparentes (“o sistema indicou”) e contra o apagamento de incertezas (quando a saída probabilística é tratada como certeza). Além disso, o próprio cenário institucional brasileiro sinaliza preocupação com auditabilidade: o Sinapses é descrito como voltado ao “treinamento, à hospedagem e à auditabilidade” de modelos de IA para uso pelo Poder Judiciário.

Responsabilidade: a cadeia decisória distribuída e a atribuição de deveres

A responsabilização por efeitos danosos, vieses ou erros em sistemas algorítmicos tende a ser dificultada pela distribuição de etapas (desenvolvimento, contratação, parametrização, implantação, uso e atualização), o que exige modelar uma “cadeia de responsabilização” em vez de buscar um único culpado. No debate institucional brasileiro, aparece como estratégia discursiva e normativa separar apoio automatizado e autoria decisória: ao apresentar usos de IA em tarefas repetitivas, afirma-se que “o magistrado... é o responsável pelas decisões e

documentos produzidos”. Essa afirmação pode ser lida como princípio de imputação (o ato jurisdicional permanece humano), mas não encerra o problema quando a recomendação, triagem ou agrupamento altera o tempo e a atenção institucional dedicada ao caso.

Do ponto de vista metodológico, Doshi-Velez e Kim sustentam que interpretabilidade é frequentemente invocada como critério de segurança, não discriminação e até como atendimento a um “right to explanation”, mas reconhecem que tais critérios “often cannot be completely quantified”, tornando a explicação uma forma de verificação qualitativa da razoabilidade do sistema. (DOSHI-VELEZ; KIM, 2017, p. 1). Para a responsabilização no setor público, isso sugere um deslocamento: não basta discutir responsabilidade “pelo resultado”; torna-se necessário discutir deveres de governança, documentação, validação, auditoria e monitoramento contínuo.

Garantias processuais e contestabilidade: contraditório, defesa e controle em ambientes tecnologicamente mediados

As garantias processuais são tensionadas quando decisões (ou etapas relevantes do procedimento) tornam-se mediadas por sistemas cuja racionalidade interna não é plenamente acessível aos afetados. Em análise crítica do caso *Wisconsin v. Loomis*, Malgieri registra que a Suprema Corte de Wisconsin considerou não haver violação do devido processo “even if they could not gain accurate information about the technicalities involved in the risk assessment process” (mesmo sem acesso a informações acuradas

sobre as técnicas do processo de avaliação de risco). (MALGIERI, 2018). Ainda que o contexto seja penal e norte-americano, o exemplo é útil para o argumento comparado: a opacidade pode ser “normalizada” institucionalmente, o que desloca o ônus de proteção para mecanismos alternativos (avisos, limites de uso, revisão humana, auditoria).

Por isso, em termos de standards, o foco tende a recair sobre (i) dever de informar que houve uso algorítmico relevante, (ii) rastreabilidade do modelo e do seu papel no caso, e (iii) meios efetivos de contestação e revisão, inclusive não judiciais. O documento do Conselho da Europa explicita esse pacote ao exigir que o uso seja “identifiable and traceable” e complementado por “effective means to contest... oversight mechanisms and effective remedies”. (CAPPELLO, 2023, p. 9). Em um desenho institucional compatível com o Judiciário brasileiro, isso conversa diretamente com o que o próprio ecossistema do Sinapses anuncia como agenda: curadoria de dados, retreino e avaliação de modelos, com transferência de conhecimento para tribunais e equipes multidisciplinares.

Conclusão

A incorporação de inteligência artificial e de automações avançadas no Judiciário brasileiro deve ser compreendida como uma mudança institucional de alta densidade: ela reorganiza fluxos de trabalho, redefine prioridades, altera práticas de triagem e padroniza rotinas, produzindo efeitos concretos sobre o modo como a jurisdição é entregue. Por isso, a questão central não é apenas aferir eficiência ou redução de acervo, mas verificar em que medida a mediação tecnológica

preserva — ou tensiona — os requisitos de legitimidade do exercício do poder estatal em um Estado de Direito.

No eixo das responsabilidades, a literatura mapeada tende a afastar soluções simplistas que buscam um único responsável e a apontar para a necessidade de uma cadeia de deveres distribuída, compatível com o ciclo de vida dos sistemas. Essa cadeia pressupõe delimitação clara de papéis (concepção, contratação, treinamento, parametrização, implantação, uso, atualização e supervisão), além de critérios verificáveis de diligência institucional, como validação antes do uso, monitoramento contínuo e protocolos de resposta a falhas. Em termos práticos, a responsabilização deixa de ser apenas um problema “ex post” e passa a exigir desenho preventivo de governança, com rastreabilidade e documentação capazes de permitir apuração e correção quando resultados indesejados ocorrerem.

No eixo da transparência e auditabilidade, a discussão converge para a ideia de que “transparência” só é juridicamente útil quando se converte em controlabilidade. Isso implica tornar o uso do sistema identificável em atos relevantes, registrar a versão e o contexto de operação, explicitar limites e incertezas e manter trilhas de auditoria que permitam reconstrução posterior. A transparência, assim, precisa ser pensada por níveis e destinatários: uma explicação compreensível ao usuário interno não é necessariamente suficiente para o controle social, e a publicidade ampla pode exigir formas intermediárias (relatórios, métricas, auditorias independentes) quando há restrições de segurança e proteção de dados. A qualidade institucional desse arranjo é medida menos pela retórica de “caixa-preta” e mais pela

existência de mecanismos consistentes de teste, revisão, correção e prestação de contas.

No eixo das garantias processuais, os achados reforçam que a mediação tecnológica reconfigura o contraditório, a ampla defesa e a motivação, sobretudo quando atua em triagens, recomendações, agrupamentos de demandas e produção de minutas. Se a parte não sabe que houve influência algorítmica relevante, ou não dispõe de instrumentos para compreender seus limites e contestar seus efeitos, cria-se uma assimetria informacional que enfraquece o controle do procedimento e pode comprometer a confiança pública no sistema de justiça. Preservar garantias, nesse contexto, exige parâmetros operacionais de contestabilidade e revisão, além de um compromisso institucional com supervisão humana efetiva — não como formalidade, mas como instância real de deliberação e correção.

Como contribuição, este estudo organiza um mapa de debates e categorias que permite ler a adoção de IA no Judiciário a partir de três lentes integradas — responsabilidade, transparência/auditabilidade e garantias — evitando tanto a tecnofilia (a crença de que a técnica resolve o problema por si) quanto o tecnopessimismo (a recusa genérica da inovação). Ao mesmo tempo, reconhece-se que o campo ainda carece de evidências empíricas comparáveis sobre desempenho, vieses, impactos distributivos e efetividade de salvaguardas, o que limita conclusões universalizantes e reforça a importância de agendas de pesquisa orientadas por dados e avaliação de impacto.

Em termos propositivos, a compatibilização entre inovação e direitos demanda um programa mínimo de governança: definição de finalidades e limites de uso; documentação e versionamento; registros de

operação e auditoria; rotinas de avaliação e retreino; canais de contestação e revisão; e critérios explícitos de suspensão quando riscos superarem benefícios. Com isso, a adoção de IA pode ser tratada não como atalho para produtividade, mas como política institucional que precisa ser justificável, controlável e compatível com a estrutura de garantias que dá sentido ao devido processo e à legitimidade da jurisdição.

Referências

CAPPELLO, M. (ed.). *Algorithmic transparency and accountability of digital services*. IRIS Special 2023-2. Strasbourg: European Audio-visual Observatory (Council of Europe), 2023. Disponível em: <https://rm.coe.int/iris-special-2023-02en/1680aeda48>. Acesso em: 28 jan. 2026.

DOSHI-VELEZ, F.; KIM, B. *Towards a rigorous science of interpretable machine learning*. arXiv:1702.08608, 2017. Disponível em: <https://arxiv.org/abs/1702.08608>; PDF: <https://arxiv.org/pdf/1702.08608.pdf>. Acesso em: 28 jan. 2026.

MALGIERI, G. Does the use of risk assessments in sentences respect the right to due process? A critical analysis of the Wisconsin v. Loomis ruling. *Law, Probability and Risk*, v. 17, n. 1, p. 45–65, 2018. Disponível em: <https://academic.oup.com/lpr/article/17/1/45/4877957>. Acesso em: 28 jan. 2026.

PHILLIPS, P. J. et al. *Four Principles of Explainable Artificial Intelligence*. NISTIR 8312. Gaithersburg: National Institute of Standards and Technology, 2021. DOI: <https://doi.org/10.6028/NIST.IR.8312>. Disponível em: <https://nvlpubs.nist.gov/nistpubs/ir/2021/nist.ir.8312.pdf>. Acesso em: 28 jan. 2026.

PNUD. *Plataforma Sinapses reúne 150 modelos de inteligência artificial*. 13 jul. 2023. Disponível em: <https://www.undp.org/pt/brazil/news/plataforma-sinapses-reune-150-modelos-de-inteligencia-artificial>. Acesso em: 28 jan. 2026.