

INTELIGÊNCIA ARTIFICIAL NA ARBITRAGEM: FERRAMENTA, ASSISTENTE OU JULGADORA?

André Guskow Cardoso

Mestre em Direito do Estado pela UFPR
Sócio da Justen, Pereira, Oliveira e Talamini
andre@justen.com.br

1. Introdução

A integração dos sistemas de inteligência artificial (IA) nos procedimentos arbitrais representa uma das transformações mais relevantes na prática da resolução de disputas nas últimas décadas. De ferramenta de pesquisa jurídica a potencial julgadora de disputas padronizadas, a IA ocupa hoje um espectro funcional amplo e crescentemente diferenciado. Compreender esse espectro é condição necessária para que árbitros, advogados e instituições arbitrais possam adotar postura informada, cautelosa e tecnicamente fundamentada diante dessas novas realidades.

O presente artigo examina as três funções que a IA pode desempenhar no contexto arbitral: (1) *ferramenta*, quando opera como instrumento de suporte técnico e operacional; (2) *assistente*, quando amplifica a capacidade cognitiva dos árbitros e das partes; e (3) *julgadora*, quando assume, total ou parcialmente, o papel de tomada de decisão. Essa estrutura tripartida permite identificar as funções específicas desempenhadas pelos sistemas de IA e os distintos níveis de salvaguarda que cada uma delas requer.

A distinção não é meramente acadêmica. Disposições genéricas sobre o uso da IA na arbitragem, seja para proibi-la, seja para autorizá-la de forma indiscriminada, podem gerar tanto restrições desproporcionais quanto exposição a riscos desnecessários. A análise funcional proposta permite otimizar o uso dessas tecnologias e minimizar seus riscos, sem abrir mão das garantias do devido processo.

2. A necessidade de uma estrutura analítica tripartida

Inteligência artificial é um termo amplo utilizado para descrever uma variedade de sistemas que frequentemente são bastante distintos entre si e desempenham funções diversas. Existem múltiplos tipos de IA e uma ampla gama de sistemas deles derivados. Por isso, o uso do termo inteligência artificial produz inevitavelmente uma generalização que tende a simplificar excessivamente as múltiplas realidades abrangidas por essa categoria de tecnologia.

Assim como não é possível generalizar o conceito de inteligência artificial nem considerar todos os sistemas de IA existentes como equivalentes, também não é possível abordar cada uso da IA na arbitragem da mesma maneira. É

necessário identificar as funções específicas desempenhadas pelos sistemas de IA, especialmente em relação aos papéis exercidos pelos árbitros, pelas partes e pelos advogados.

Essa estrutura tripartida baseia-se em tipologias funcionais desenvolvidas na literatura sobre IA (por exemplo, Jieqiong Cao et al., Not All Forms of Artificial Intelligence Are Perceived Equal: AI Functions and Work Outcomes. *Journal of Open Innovation: Technology, Market, and Complexity*, 11(2), 2025. Disponível em: <https://doi.org/10.1016/j.joitmc.2025.100521>), que distinguem entre sistemas assistivos, aumentados e autônomos, bem como em análises de sistemas jurídicos de IA que enfatizam a diversidade de funções e riscos. No campo da arbitragem, diversas contribuições distinguem igualmente entre o uso da IA como ferramenta de apoio procedimental e propostas em que a IA auxilia, ou potencialmente substitui, o árbitro na função adjudicatória. A classificação aqui adotada (*ferramenta, assistente e julgadora*) sistematiza esses debates em um único referencial analítico.

Essas categorias refletem diferentes graus de interação tecnológica, exigindo níveis diferenciados de salvaguardas e regulação. Enquanto o uso da IA como ferramenta visa principalmente à eficiência, seu uso como assistente ou como julgadora exige salvaguardas muito mais rigorosas, a fim de preservar a integridade do processo jurídico. Reconhecer essa diferença constitui pressuposto indispensável de qualquer esforço regulatório sério sobre o tema.

3. A IA como ferramenta

Em sua função mais elementar, os sistemas de IA servem a fins eminentemente técnicos: gestão e organização de documentos, tradução de textos, agendamento de audiências, visualização de dados processuais, identificação de prazos e automação de tarefas administrativas repetitivas. Aumentam a eficiência sem afetar o raciocínio adjudicatório. Nesse sentido, a IA representa uma evolução funcional comparável à substituição das máquinas de escrever pelos computadores – uma mudança nos instrumentos de trabalho, não na essência do trabalho em si.

O uso da IA como ferramenta é o mais amplamente aceito no meio arbitral. A pesquisa conduzida em 2025 pelo escritório White & Case em colaboração com a Queen Mary University of London (<https://www.qmul.ac.uk/arbitration/media/arbitration/docs/White-Case-QMUL-2025-International-Arbitration-Survey-report.pdf>) confirmou que os principais usos atuais da IA na arbitragem incluem pesquisa fática e jurídica (relatada por 64% dos respondentes), análise de dados e revisão de documentos – todos compatíveis com a função de ferramenta. O uso nessa categoria é objeto de aprovação praticamente unânime entre os profissionais consultados na pesquisa.

Nessa modalidade de uso, os riscos são mais restritos e gerenciáveis: a possibilidade de vazamento de dados processuais sensíveis e a propagação de vieses existentes nos dados de treinamento. As salvaguardas correspondentes são relativamente menos onerosas – controles de acesso, protocolos de

anonimização de dados e validação independente dos sistemas – e podem ser facilmente incorporadas a regras processuais existentes.

4. A IA como Assistente

A IA como assistente apresenta autonomia aumentada, ainda que limitada. Nessa função, a IA vai além do suporte operacional para interagir diretamente com o conteúdo substantivo do procedimento arbitral: redige seções de laudos, sugere estruturas de fundamentação, identifica precedentes relevantes, resume petições volumosas ou extrai os argumentos centrais de documentos extensos. Seu papel é ampliar a capacidade cognitiva humana (o que se pode denominar "cognição aumentada"), não substituí-la.

A pesquisa da Queen Mary University avaliou a percepção de adequação do uso da IA por árbitros em diferentes tarefas. A grande maioria dos respondentes (72%) relatou nunca utilizar IA para redigir petições, citando preocupações com precisão e raciocínio. Há ampla aprovação para usos assistivos em tarefas com menor grau de discricionariedade, como assistência ao cálculo de danos, custos e juros (77% de aprovação) e sumarização de petições ou provas (66%). Em contrapartida, há resistência ao uso da IA em tarefas que exigem o exercício de discricionariedade e julgamento: apenas 23% dos respondentes aprovam o uso da IA para redigir as seções de fundamentação jurídica de sentenças ou decisões. Para os próximos cinco anos, 52% esperam que os árbitros passem a depender crescentemente da IA, e 48% acreditam que os procedimentos se tornarão mais rápidos.

O risco central associado ao uso assistivo da IA é o chamado "viés de automação" – a tendência humana a aceitar as sugestões de sistemas automatizados sem exercer o escrutínio crítico adequado. Sophie Nappert identifica o "fator uau" como manifestação concreta desse risco: um viés cognitivo-emocional que surge quando nos deparamos com resultados que superam nossa própria capacidade, especialmente quando provêm de agentes não humanos (The "WOW" Factor: Human Cognition in the Age of Algorithmic Summarisation).

<https://www.arbtech.io/blog/the-wow-factor-human-cognition-in-the-age-of-algorithmic-summarisation>

Segundo a autora, o "fator uau" refere-se a "um viés cognitivo-emocional que surge quando nos deparamos com resultados que superam nossa própria capacidade – especialmente quando provêm de agentes não humanos, como a IA. Há uma pesquisa crescente demonstrando que os seres humanos têm uma tendência padrão a atribuir autoridade e até superioridade epistêmica a máquinas que apresentam informações com alta fluência e velocidade.

A supervisão humana efetiva sobre os sistemas assistivos de IA requer salvaguardas procedimentais estruturadas: dever explícito de raciocinar de forma independente, justificativa documentada quando se recorre à análise assistida por IA, transparência quanto ao papel das ferramentas e orientação institucional voltada à mitigação do viés de automação. A delegação dessas funções a secretários de tribunais arbitrais não exige o árbitro de entender, em

termos gerais, como os sistemas operam e como podem introduzir vieses nocivos, pois essas são questões que afetam diretamente o cerne da tomada de decisão – função primordial e inalienável do árbitro.

5. A IA como julgadora

Em contextos limitados e estritamente definidos, como litígios decorrentes de contratos inteligentes (*smart contracts*) ou arbitragem *online* de baixo valor, os sistemas de IA podem ser capazes de realizar formas de adjudicação baseada em regras, tipicamente caracterizadas por insumos e dados padronizados, regras claramente definidas e discricionariedade limitada (Cf., por exemplo, P. Han, AI-Powered Digital Arbitration Framework Leveraging Smart Contracts and Electronic Evidence Authentication. 15 Scientific Reports 37327 (2025). <https://doi.org/10.1038/s41598-025-21313-x>; e Luís de Lima Pinheiro, Arbitration, Blockchain and Artificial Intelligence, 28 mai. 2025. <https://ssrn.com/abstract=5272787>).

Nessas situações, a IA opera essencialmente como um árbitro algorítmico: dada uma entrada (o conjunto de fatos e as cláusulas contratuais relevantes), o sistema produz uma saída (a decisão) de acordo com parâmetros predefinidos, sem necessidade de julgamento de valor.

Em contrapartida, os litígios complexos envolvem padrões jurídicos com textura aberta, avaliação de provas, julgamentos de credibilidade e ponderação normativa – aspectos que não são compatíveis com a tomada de decisão completamente autônoma pela IA. A avaliação da credibilidade de uma testemunha, a ponderação entre princípios jurídicos contraditórios, a interpretação de cláusulas contratuais ambíguas à luz da intenção das partes: todas essas operações exigem um tipo de julgamento específico, contextual e eticamente orientado que os sistemas de IA atuais não são capazes de exercer de forma autônoma e confiável.

Embora a substituição completa de julgadores humanos em casos complexos permaneça improvável no futuro previsível, os desenvolvimentos atuais apontam para um aumento gradual da participação algorítmica em funções determinativas, especialmente quando envolverem tarefas estruturadas e repetitivas. Nesse cenário, cabe questionar não se a IA algum dia poderá substituir o árbitro humano em toda a sua extensão funcional, mas em que medida e em que condições ela pode assumir funções determinativas parciais, especialmente em segmentos do processo decisório que sejam suficientemente estruturados e em que a participação algorítmica possa ser adequadamente auditada, supervisionada e contestada pelas partes.

Nesse sentido, o Regulamento de Inteligência Artificial da União Europeia (AI Act, Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho, de 13 de junho de 2024) classifica como de alto risco os sistemas de IA que influenciam materialmente resultados em áreas sensíveis como emprego, educação, aplicação da lei e justiça (art. 6º, nº 2, e Anexo III), impondo obrigações rigorosas de gestão de riscos, governança de dados, transparência e avaliações de conformidade – salvo quando se trate de tarefas

procedimentais restritas ou preparatórias que não influenciem materialmente as decisões (art. 6º, nº 3).

Os sistemas de IA de alto risco estão sujeitos a requisitos rigorosos, incluindo a implementação de um sistema de gestão de riscos (art. 9º), a garantia de alta qualidade na governança de dados (art. 10), a manutenção de documentação técnica (art. 11), obrigações de transparência (art. 13), garantias de precisão, robustez e cibersegurança (art. 15) e a realização de avaliações de conformidade (art. 43). Além disso, os fornecedores são obrigados a estabelecer um sistema de gestão da qualidade (art. 17), manter registros (art. 19), emitir declarações de conformidade da UE (art. 47), afixar a marcação CE, registrar os sistemas de IA e adotar medidas corretivas quando necessário (art. 16).

Essa estrutura normativa, embora voltada ao contexto europeu e a sistemas judiciais ou decisórios públicos, oferece referências relevantes para a governança da IA na arbitragem.

A ausência de supervisão humana significativa – pela qual o árbitro meramente "referenda" a recomendação de um sistema de IA sem exame crítico genuíno – pode ser caracterizada como uma abdicação do mandato adjudicatório, com potenciais consequências para a futura exequibilidade da sentença arbitral. Se um tribunal arbitral se apoiar em ferramentas de IA opacas, do tipo "caixa-preta", a parte vencida pode arguir que foi privada de um julgamento justo – seja pela falta de explicabilidade, seja pelo eventual viés algorítmico dos dados de treinamento.

6. Comparação das Três Funções

As três funções da inteligência artificial na arbitragem podem ser estruturadas no seguinte quadro analítico comparativo:

Papel do sistema de IA	Descrição & função	Exemplos práticos	Objetivo principal
(1) Ferramenta	A IA é utilizada para realizar tarefas mecânicas ou fornecer suporte técnico e operacional.	Pesquisa de jurisprudência, tradução, revisão/análise de documentos e gestão de prazos.	Aumentar a eficiência do fluxo de trabalho, sem influenciar o conteúdo da decisão.
(2) Assistente	Atua como copiloto para o árbitro ou as partes, com autonomia limitada.	Propor minutas de sentença arbitral ou relatórios, sugerindo linhas de raciocínio e auxiliando na análise das provas.	Ampliar a capacidade cognitiva humana ("cognição aumentada").

Papel do sistema de IA	Descrição & função	Exemplos práticos	Objetivo principal
(3) Julgadora	A IA assume o papel de tomada de decisão, de forma total ou parcial.	Arbitragem automatizada em disputas padronizadas ou parametrizadas (ex. smart contracts).	Decidir disputas com base em parâmetros estritamente predefinidos.

A analogia com os níveis de automação veicular é útil para visualizar as distinções. Como ferramenta, a IA opera como um GPS que orienta o itinerário – o motorista mantém o controle total. Como assistente, assemelha-se a um piloto automático sob supervisão humana – o sistema executa funções relevantes, mas o motorista deve estar apto a intervir a qualquer momento. Como julgadora, funciona como um veículo totalmente autônomo – o sistema decide, executa e encerra sem intervenção humana ao longo do percurso.

Cada degrau na escala de autonomia implica um salto qualitativo nos riscos e nas exigências de governança. É precisamente por isso que a distinção funcional tem implicações normativas, éticas e procedimentais de primeira ordem para a integridade do processo arbitral.

7. IA na Arbitragem: Riscos, Regulação e Governança

Levando em conta as três funções distintas, é possível identificar distintos níveis de salvaguardas: à medida que aumenta o grau de autonomia de um sistema, deve aumentar também a intensidade dos controles a ele impostos. Essa abordagem é análoga à regulação baseada em risco do AI Act, embora as diferentes funções devam desencadear respostas distintas.

O quadro abaixo apresenta os riscos primários associados a cada função e as salvaguardas correspondentes:

Função	Riscos Primários	Salvaguardas Exigidas
Ferramenta	Vazamento de dados; propagação de viés existente nos dados de treinamento.	Controles de acesso; protocolos de anonimização; validação independente.
Assistente	Dependência excessiva; viés de automação; viés cognitivo ("fator uau").	Supervisão humana integrada; dever de raciocínio independente; transparência sobre uso; mitigação do viés de automação.
Julgadora	Crise de legitimidade; opacidade ("caixa-preta"); falta de	Garantia de explicabilidade; mecanismos de recorribilidade; certificação independente de sistemas;

Função	Riscos Primários	Salvaguardas Exigidas
	responsabilização; impugnabilidade da sentença.	divulgação obrigatória às partes.

Na arbitragem institucional, como no caso da AAA-ICDR, da CCI ou da LCIA, câmaras arbitrais podem estabelecer salvaguardas por meio de suas regras processuais, orientações específicas ou requisitos de divulgação. A adoção institucionalizada de uma plataforma dedicada e habilitada por IA para a administração de procedimentos arbitrais permite um grau muito maior de controle e transparência do que um modelo excessivamente aberto, em que cada árbitro e parte usa qualquer ferramenta tecnológica de sua livre escolha – modelo que cria complexidades adicionais e pode, em última análise, minar a transparência do procedimento e dificultar a avaliação dos efeitos reais da IA sobre a decisão arbitral.

Na arbitragem *ad hoc*, a adoção de salvaguardas comparáveis é consideravelmente mais difícil. A ausência de uma instituição que possa impor regras uniformes torna a definição de parâmetros de governança – e, sobretudo, seu cumprimento – um desafio considerável. Nesse contexto, a adoção voluntária de protocolos e diretrizes por árbitros e partes parece representar a via mais promissora no curto prazo.

8. Indo além da compreensão convencional

Reconhecer que os sistemas de IA desempenham funções distintas, implicam diferentes níveis de risco e requerem salvaguardas específicas é essencial ao se buscar regular o uso da IA na arbitragem. Disposições genéricas podem acarretar restrições desproporcionais a sistemas utilizados meramente como ferramentas; ao mesmo tempo, considerar a função específica permite otimizar o uso e minimizar os riscos associados. Note-se que essa análise não aborda restrições gerais aplicáveis a todos os sistemas de IA, como os riscos de alucinações ou as limitações decorrentes de vieses nos dados de treinamento – preocupações que devem igualmente ser avaliadas à luz da importância da tarefa e dos riscos envolvidos.

A revisão sistemática conduzida por Rodríguez-Salcedo e outros autores sobre a integração da IA nos Métodos Alternativos de Resolução de Disputas (MARD), que sintetizou 30 artigos publicados entre 2010 e 2025, revelou que a IA melhora a acessibilidade e o desempenho operacional dos MARD principalmente por meio da automação de tarefas administrativas e da assistência na análise de documentos, mas que sua implementação coloca desafios não resolvidos: vieses algorítmico, confidencialidade de dados e ausência de marcos regulatórios adequados (Rethinking Mediation and Arbitration in the Age of Artificial Intelligence: A Systematic Review. Salud, Ciencia y Tecnología, 15 out. 2025).

<https://sct.ageditor.ar/index.php/sct/article/view/2316>;
<https://doi.org/10.56294/saludcyt20252316>).

Os profissionais tendem a rejeitar o uso da IA nas decisões finais críticas, preferindo empregá-la como ferramenta complementar de apoio.

O caminho mais sustentável, segundo os autores, reside no conceito de "justiça aumentada" (*augmented justice*), em que a IA apoia a *expertise* humana ao aumentar a eficiência procedimental, preservando a supervisão humana. Embora a IA fortaleça o desempenho operacional dos MARD, ela não pode substituir o julgamento ético ou a empatia inerentes à participação humana na administração da justiça.

Nesse sentido, a noção de **IA constitucional** oferece uma perspectiva complementar. Trata-se de uma técnica de alinhamento de modelos de linguagem que incorpora um conjunto de princípios normativos — uma espécie de "constituição" — diretamente ao processo de treinamento, orientando o comportamento do sistema em conformidade com valores previamente definidos (Sophie Nappert, Benjamin Malek e Fernanda Carvalho Dias de Oliveira Silva, What is Constitutional AI and Why Does it Matter for International Arbitration? <https://legalblogs.wolterskluwer.com/arbitration-blog/>. Acesso em: 12 dez. 2025).

Sophie Nappert, Benjamin Malek e Fernanda Carvalho Dias de Oliveira Silva explicam que a IA constitucional "substitui grande parte da supervisão humana exigida durante o treinamento por um conjunto escrito de princípios normativos – a 'constituição' –, que a IA utiliza para orientar seu comportamento. Esses princípios, explicados em linguagem simples (como 'evitar conteúdo prejudicial' ou 'ser respeitoso'), funcionam como salvaguardas internas".

O termo foi introduzido pela empresa Anthropic em 2022, que definiu previamente um conjunto escrito de princípios para que o próprio modelo avalie e corrija suas respostas (Yuntao Bai, Saurav Kadavath, Sandipan Kundu et al., Constitutional AI: Harmlessness from AI Feedback, 2022. <https://arxiv.org/abs/2212.08073>). A Anthropic desenvolveu também a noção de IA Constitucional Coletiva, em que o conjunto de princípios é construído com a participação do público ou de múltiplas partes interessadas, de modo a refletir valores sociais mais amplos.

No contexto da arbitragem, a IA constitucional poderia internalizar princípios como imparcialidade, neutralidade, transparência e separação entre fatos e julgamentos de valor – por exemplo, por meio de resumos de provas que evitem atribuir credibilidade, "constituições" definidas pelas partes e modelos treinados com base nas regras das instituições arbitrais (CCI, LCIA, UNCITRAL). As principais vantagens incluem maior interpretabilidade e previsibilidade do sistema. Já os principais desafios envolvem a definição participativa da constituição e a necessidade contínua de julgamento humano nas decisões jurídicas substantivas.

Há ainda a dimensão dos **agentes de IA** – sistemas autônomos e orientados a objetivos específicos, capazes de executar tarefas e comunicar-se com fontes de dados externas. Ao contrário de um *chatbot* convencional, que funciona como um "copiloto" que gera conteúdo, um agente opera como um "piloto automático", capaz de usar ferramentas externas para executar fluxos de trabalho complexos de forma autônoma (Joe Kwon, AI Agents: Governing Autonomy in the Digital Age, 2025 https://cdn.prod.website-files.com/65af2088cac9fb1fb621091f/682f96d6b3bd5a3e1852a16a_AI_Agents_Report.pdf).

A autonomia dos agentes de IA coloca desafios adicionais para a estrutura tripartida aqui proposta, pois esses sistemas podem operar em múltiplos papéis simultaneamente, tornando mais difícil a delimitação das salvaguardas aplicáveis.

9. Conclusão

A integração da inteligência artificial na arbitragem não é uma questão do futuro: é uma realidade presente e crescente. A questão decisiva não é se a IA será utilizada, mas como ela pode ser regulada e governada de maneira consistente com a equidade, a transparência e a dignidade humana que fundamentam o processo arbitral.

A estrutura tripartida proposta – IA como *ferramenta*, como *assistente* e como *julgadora* – oferece um referencial analítico que permite identificar, de forma precisa e proporcional, os riscos associados a cada modalidade de uso e as salvaguardas que cada uma delas requer. Essa distinção é o pressuposto indispensável de qualquer esforço regulatório sério: sem ela, corre-se o risco de proibir o que é benéfico e autorizar o que é perigoso.

Enquanto o uso da IA como ferramenta é amplamente aceito e requer salvaguardas relativamente simples, o uso como assistente exige supervisão humana estruturada e letramento tecnológico adequado. O uso como julgadora, por sua vez, só é viável em contextos limitados e estritamente definidos, com requisitos elevados de explicabilidade, recorribilidade e certificação independente. Em todos os casos, a responsabilidade última pelo processo arbitral e por suas decisões deve permanecer com o árbitro humano.

O árbitro do futuro deve estar equipado com o necessário letramento tecnológico – não apenas para usar essas ferramentas, mas também para questioná-las, interpretá-las e limitar sua influência quando necessário (nesse sentido, Sophie Nappert, The Impact of Artificial Intelligence on the Duties, Qualifications and Selection of International Arbitrators, Oxford Law Pro, Thought Leadership series, 2025, <https://academic.oup.com/info-oxford-law-pro/pages/artificial-intelligence-arbitrators>).

A IA não substituirá os árbitros; ao contrário, redefinirá o próprio significado da adjudicação e do processo decisório, transformando a arbitragem em um

ambiente de "justiça aumentada" em que a assistência tecnológica não suplanta a responsabilidade ética humana.

Informação bibliográfica do texto:

CARDOSO, André Guskow. Inteligência Artificial na Arbitragem: Ferramenta, Assistente ou Julgadora? *Informativo Justen, Pereira, Oliveira e Talamini*, Curitiba, n.º 230, abril de 2026, disponível em <https://justen.com.br/>, acesso em [data].