

Generative Artificial Intelligence and Legal Decision-making

André Guskow Cardoso^{*}, Elizabeth Chan^{**}, Luísa Quintão^{***} & Cesar Pereira^{****}

This article explores the transformative impact of generative artificial intelligence (GenAI) on legal decision-making processes. It also examines disclosure obligations and the challenges AI-assisted decisions pose in litigation and arbitration.

A case study using ChatGPT as arbitrators in the Willem C. Vis International Commercial Arbitration Moot provides practical insights into AI's potential and limitations in arbitral decision-making. The experiment highlights key learnings about technology constraints, due process concerns, and the enforceability of AI-assisted decisions.

The article concludes with reflections on integrating AI into legal frameworks, emphasizing the need for updated regulations and best practices to ensure transparency, fairness, and accountability in AI applications within the legal domain.

Keywords: Generative AI, Legal decision-making, AI-assisted arbitration, Disclosure obligations, Moot court, Legal technology, Ethical AI

I INTRODUCTION

The transformative impact of generative artificial intelligence (**GenAI**) on legal decision-making processes is increasingly evident. This article delves into AI's role in the legal domain. It examines its potential to revolutionize judicial and arbitration decision-making processes while addressing the significant challenges and limitations accompanying its adoption.

The first section (II) explores AI-assisted legal decision-making, highlighting recent advancements and implementations globally. We examine notable experiments and applications, such as a Pakistani court's use of ChatGPT-4, AI integration in China's judicial system, and experiments with Claude 3 Opus in adjudicating US Supreme Court cases. This section sets the stage by demonstrating AI's potential to enhance efficiency, accuracy, and accessibility in legal decision-making.

The second section (III) addresses the limitations of using AI in legal decision-making. We discuss common concerns such as biases, the 'black box' phenomenon, hallucinations, and the specific challenges in domain-specific AI systems. This section also covers the ethical issues

surrounding AI use, including the difficulty of translating ethical standards into code and the necessity of maintaining human oversight.

The third section (IV) focuses on disclosure obligations for decision-makers, such as judges and arbitrators, who use AI tools. We examine guidelines and frameworks from various jurisdictions, including New Zealand, the United Kingdom, the United States, the European Union, and Brazil. This section emphasizes the importance of transparency, informed consent, and accountability in using AI in legal settings.

In the fourth section (V), we analyse AI-assisted decision-making challenges. We discuss the procedural and substantive issues that may arise and consider how these challenges mirror those related to the involvement of tribunal secretaries. This section underscores the need for clear guidelines, transparency, and informed consent to ensure the arbitration process remains fair and accountable.

The fifth section (VI) presents a case study involving the use of ChatGPT as arbitrators in the Willem C. Vis International Commercial Arbitration Moot. This practical experiment provides insights into the potential and

Notes

^{*} Partner at Justen, Pereira, Oliveira & Talamini, based in Curitiba. He is an active member of ArbTech, a Tech & ADR think-tank. Email: andre@justen.com.br.

^{**} Foreign Legal Manager at Stevenson, Wong & Co., based in Hong Kong. Email: elizabeth.chan@sw-hk.com.

^{***} Partner at Justen, Pereira, Oliveira & Talamini, based in Sao Paulo. He is a Chartered Arbitrator (C.Arb) and Fellow of the Chartered Institute of Arbitrators (FCIarb). Email: cesar@justen.com.br.

^{****} Associate at Justen, Pereira, Oliveira & Talamini, based in Sao Paulo. Email: luisa.quintao@justen.com.br. The authors are very grateful to Lorenzo Galan Miranda for his research assistance.

limitations of AI in arbitration, highlighting key learnings about technology constraints, due process concerns, and the enforceability of AI-assisted decisions.

Finally, the article concludes with reflections on integrating AI into legal frameworks. We emphasize the need for updated regulations and best practices to ensure transparency, fairness, and accountability in AI applications within the legal domain. The journey towards trustworthy and accountable AI in legal decision-making is ongoing and requires continuous reflection, adaptation, and innovation.

2 AI-ASSISTED LEGAL DECISION-MAKING

AI tools have long had the potential for judicial and arbitration use.¹ As Richard Susskind explains, the evolution of AI in law began in the 1960s.² Early AI systems, primarily machine learning-based, were used for document and information processing but required structured data. These systems were not widely adopted due to the lack of structured information in judicial and arbitration proceedings. Data from these proceedings are ‘unstructured’ because they do not follow a consistent pattern, as each court and arbitral tribunal has its unique style and organization for awards and sentences. Variations in structure, style, order of presentation, and terminology are common when dealing with judicial and arbitral awards and proceedings.³

GenAI builds on these older AI systems rather than replacing them, marking a continuous advancement in AI techniques within the legal field.⁴ Large Language Models (LLMs) have revolutionized the use of AI in law practice and decision-making by enabling more advanced and nuanced legal information processing.

This section is divided into three subsections. The first subsection examines recent advancements in AI use in courts globally, including notable experiments and the development of explainable AI tools. The second subsection discusses the application of AI in alternative dispute resolution, highlighting examples from arbitration and

innovative approaches to non-binding neutral evaluations. The third subsection explores AI’s role in regulatory and administrative contexts, detailing tools used in public procurement and regulatory decision-making to enhance transparency and combat corruption.

2.1 AI-assisted Legal Decision-Making in the Courts

Recent developments in AI-assisted legal decision-making illustrate a transformative shift in how court processes and judgments are approached globally.

Pakistan: One notable experiment by a Pakistani Court involved using ChatGPT-4 to identify the issues for deciding an appeal of an application for a temporary injunction.⁵ The Court asked Chat GPT-4 for the principles for granting an injunction in a civil case in Pakistan.⁶ The Court found Chat GPT-4’s response ‘impressive’ and ‘helpful in crafting a quality order or judgment’.⁷

China: In China, AI has been integrated into the judicial system, with developments including the establishment of the country’s first ‘smart court’ in Hangzhou in 2017 and the deployment of robot judges and guide robots like ‘Yun Fan’ in Xinyang and ‘Xiao Chong’ in Chongzhou.⁸ These robots assist with various legal processes, from answering questions based on millions of legal cases to providing litigants with pre-trial predictions and legal consultations.⁹ Further, AI judges, such as ‘Xiaozi’, have successfully adjudicated small claims and dispute cases under human supervision, demonstrating potential for unbiased and efficient decision-making in straightforward legal domains.¹⁰

United States: In a compelling experiment, Adam Unikowsky explored the capabilities of Claude 3 Opus, an advanced AI, in adjudicating Supreme Court cases.¹¹ By inputting briefs from every Supreme Court merits case decided so far in a specific term into Claude, Unikowsky found that the AI performed impressively, aligning its

Notes

¹ Cesar Pereira, André Guskow Cardoso, Christine Santini & Leonardo Souza-McMurtrie, *AI and Dispute Resolution: a Powerful (New) Tool in International Commercial Arbitration*, in *International Commercial Arbitration Practice: 21st Century Perspectives* (Horacio Grigera Naon, Paul Eric Mason, & Gustavo Moser eds, LexisNexis, Matthew Bender 2023).

² *Ibid.*

³ Currently, there are initiatives as SALI – Standards Advancement for the Legal Industry, that intend to define broad standards for legal data. See www.sali.org (accessed 20 Jul. 2024).

⁴ *Ibid.*

⁵ Muhammad Iqbal s/o Muhammad Din, Caste Choghatta, r/o Hussain Park Lahore v. Zayad (deceased) son of Muhammad Aalim through legal heirs, ETC., Civil Appeal No. 11 of 2023, 28 Mar. 2023, para. 18.

⁶ *Ibid.*, para. 19.

⁷ *Ibid.*, para. 21.

⁸ N. Wang, M. Y. Tian, *Intelligent Justice: AI Implementations in China’s Legal Systems*, in *Artificial Intelligence and Its Discontents. Social and Cultural Studies of Robots and AI* 211–212 (A. Hanemaayer ed., Palgrave Macmillan, Cham 2022), doi: 10.1007/978-3-030-88615-8_10 (accessed 8 Jul. 2024).

⁹ *Ibid.*

¹⁰ *Ibid.*

¹¹ Adam Unikowsky, *In AI we Trust, part II*, Adam’s Legal Newsletter (16 Jun. 2024), <https://adamunikowsky.substack.com/p/in-ai-we-trust-part-ii> (accessed 8 Jul. 2024).

decisions with those of the Court in 27 out of 37 cases. Even when Claude's decisions differed, Unikowsky considered them reasonable and well-argued. Unikowsky noted that Claude's performance as a law clerk was comparable to, if not better than, human clerks in terms of accuracy and creativity, with the added benefit of unparalleled efficiency. This experiment suggests that while AI may not yet replace judges, it can significantly enhance the judicial process by supporting clerks with insightful and rapid analyses.

United States: Judge Kevin Newsom of the Eleventh Circuit, in his concurring opinion in the case of *Snell v. United Specialty Insurance Company*, provided an intriguing insight into the use of LLMs to assist in his interpretive process.¹² Judge Newsom wrote a separate opinion discussing his use of ChatGPT to assist in understanding the ordinary meaning of 'landscaping'. Specifically, Judge Newsom detailed how he used ChatGPT to get an additional perspective on whether the installation of an in-ground trampoline fell under the term 'landscaping' as used in the insurance policy under scrutiny.¹³

Judge Newsom found the responses from ChatGPT helpful. When asked to define 'landscaping', ChatGPT provided a broad definition that included altering the visible features of an area of land for aesthetic or practical purposes, which resonated with Newsom's initial thoughts.¹⁴

However, Judge Newsom also noted the limitations and potential drawbacks of relying on LLMs. He acknowledged the 'hallucinations' of LLMs, where AI generates inaccurate or fabricated information, and highlighted the concern that LLMs might not fully capture offline speech patterns, potentially overlooking the linguistic contributions of underrepresented communities.¹⁵

Despite these concerns, Judge Newsom saw value in the potential of LLMs to assist in legal interpretation. He

suggested that, with further refinement and transparency in their use, LLMs could complement traditional tools like dictionaries and canons of construction, offering a more democratized and accessible means of understanding ordinary language in legal contexts.¹⁶ He emphasized the need for caution and humility in integrating these technologies, reflecting a balanced view of their promising capabilities and inherent limitations.¹⁷

Europe: Further, the development and integration of explainable AI (XAI) tools in legal reasoning underscore efforts to make AI-assisted decisions more transparent and understandable to parties and their counsel.¹⁸ An example of XAI in the legal context is a programme developed by researchers at the Universities of Manchester and Liverpool intended to recommend to laypersons whether a case they wanted to submit to the European Court of Human Rights would be deemed admissible.¹⁹ The programme would give reasons for its recommendation.²⁰

2.2 AI-assisted Legal Decision-Making in Alternative Dispute Resolution

AI has moved into arbitration, too, with the Guangzhou Arbitration Commission reporting in 2023 that its AI arbitration assistant had resolved a dispute between two domestic private companies.²¹ These examples signal a move towards greater acceptance of AI to assist in resolving disputes, although usually under human supervision. This trend is especially illustrated by the publication of the Silicon Valley Arbitration and Mediation Center's (SVAMC) Guidelines on the Use of AI in Arbitration (SVAMC Guidelines) in April 2024 (discussed below).²² The SVAMC Guidelines are best practice guidelines for responsible AI use in arbitration proceedings.²³

Sapna Jhangiani KC suggests an innovative approach to enhancing dispute resolution through the use of AI in non-binding neutral evaluation.²⁴ This method, as

Notes

¹² *Snell v. United Specialty Insurance Company* (22–12581, USCA11) (28 May 2024), <https://media.ca11.uscourts.gov/opinions/pub/files/202212581.pdf> (accessed 28 Jul. 2024).

¹³ *Ibid.*, at 9.

¹⁴ *Ibid.*

¹⁵ *Ibid.*, at 21–22.

¹⁶ *Ibid.*, at 25.

¹⁷ *Ibid.*

¹⁸ Joe Collenette, Katie Atkinson & Trevor Bench-Capon, *Explainable AI tools for legal reasoning about cases: A study on the European Court of Human Rights*, 317 *Artificial Intelligence* (2023), doi: 10.1016/j.artint.2023.103861 (accessed 8 Jul. 2024).

¹⁹ *Ibid.*, at 10.

²⁰ *Ibid.*, at 11 and 12.

²¹ Qiu Quanlin, *AI arbitration used for dispute in Guangzhou* (China Daily 1 Sep. 2023), www.chinadaily.com.cn/a/202309/01/WS64f13406a310d2dce4bb34ad.html (accessed 8 Jul. 2024).

²² Silicon Valley Arbitration and Mediation Center, *SVAMC Publishes Guidelines on the Use of Artificial Intelligence in Arbitration* (30 Apr. 2024), <https://svamc.org/svamc-publishes-guidelines-on-the-use-of-artificial-intelligence-in-arbitration> (accessed 8 Jul. 2024).

²³ An author of this article, Elizabeth Chan, served on the nine-member drafting committee.

²⁴ Sapna Shangiani KC, 'Ambulances, Cliffs, and the Extinction of Arbitration Practitioners', *Kluwer Arbitration Blog* (23 Mar. 2024), https://arbitrationblog.kluwerarbitration.com/2024/03/23/ambulances-cliffs-and-the-extinction-of-arbitration-practitioners/?utm_source=pocket_saves (accessed 8 Jul. 2024).

highlighted in the ICC Commission Guide on Effective Conflict Management, involves a third-party neutral providing an evaluation of the parties' differences to facilitate settlement.²⁵ The evaluation is based on limited evidence, making it a cost-effective alternative to a full hearing.

Jhangiani proposes using AI tools to generate these neutral evaluations rapidly and at minimal cost by analysing previous decisions to predict likely outcomes.²⁶ Despite the absence of detailed reasoning in AI-generated evaluations, they hold significant value by offering predictive insights that can guide parties towards settlement. This AI application could increase the use of neutral evaluation, making it a more attractive option in the toolkit of dispute resolution mechanisms.

2.3 AI-assisted Legal Decision-Making in the Administrative and Regulatory Context

The exploration of AI in regulatory decision-making and the discussions around its broader implications highlight the technology's applicability beyond judicial decisions, extending into regulatory and administrative law.

For example, the Danish Competition and Consumer Authority has developed a tool named Bid Viewer, designed to scrutinize public procurement bids for signs of potential cartel behaviour.²⁷ This innovative tool is capable of identifying instances where companies might have divided geographical territories to avoid bidding against each other, detecting patterns of firms alternately submitting the lowest bids, and spotting situations where two organizations consistently refrain from competing against each other in procurement processes.²⁸

In Brazil, the Federal Court of Accounts (TCU) is leveraging AI and data analysis to enhance regulatory decision-making and combat corruption.²⁹ Like Denmark's Bid Viewer, the TCU's advanced tools, such as LabContas and LabCor, utilize extensive public administration databases to identify risks and vulnerabilities. AI-driven systems like Alice (Public Tender, Contract and Bid Analyser) and SAO (Budget Analysis System)

automate the analysis of public tenders and budgets, promoting transparency and efficiency. By integrating AI into its auditing processes, the TCU can generate semi-automated reports and data-driven insights, streamlining audits and improving the effectiveness of monitoring public resources. This innovative use of AI in the regulatory space underscores the technology's potential to foster accountability and enhance governance.

3 LIMITATIONS IN USING AI IN LEGAL DECISION-MAKING

Despite the potential benefits of using AI tools in legal decision-making, there are notable limitations.³⁰ This section is divided into three subsections to explore these constraints comprehensively.

The first subsection discusses common concerns such as biases, the 'black box' phenomenon, and hallucinations, explaining how these issues arise and their implications for AI reliability and transparency.

The second subsection delves into the challenges involving AI systems in legal contexts, including the confidential nature of arbitration data, the risk of perpetuating human biases, and the tension between prediction and legal interpretation.

The third subsection addresses the ethical concerns surrounding AI use and highlights the importance of integrating ethical principles into AI systems to ensure fairness, transparency, and accountability.

3.1 General Limitations in Using AI

Common concerns about AI include biases, the 'black box' phenomenon, and hallucinations.

Biases: Biases can stem from training datasets that reflect societal biases or from the developers themselves. Even the choice of training data can introduce bias.³¹ This can lead to systems that perpetuate existing inequalities and produce unfair or inaccurate outcomes.

Notes

²⁵ ICC Commission on Arbitration and ADR, *Effective Conflict Management* (2023), para. 90, <https://iccwbo.org/wp-content/uploads/sites/3/2023/07/2023-ICC-Effective-Conflict-Management.pdf> (accessed 8 Jul. 2024).

²⁶ Sapna Shangiani KC, *Ambulances, Cliffs, and the Extinction of Arbitration Practitioners*, Kluwer Arbitration Blog (23 Mar. 2024), <https://arbitrationblog.kluwerarbitration.com/2024/03/23/ambulances-cliffs-and-the-extinction-of-arbitration-practitioners> (accessed 8 Jul. 2024).

²⁷ HSF, *AI in regulatory Decision making – is it the end of the World?*, HSF Notes (27 Sep. 2023), <https://hsfnotes.com/publiclaw/2023/09/27/ai-in-regulatory-decision-making-is-it-the-end-of-the-world> (accessed 8 Jul. 2024).

²⁸ *Ibid.*

²⁹ Tribunal de Contas da União, *TCU invests in Institutional Partnerships, data Analysis and the use of AI in the Fight Against Corruption*, Notícias TCU (3 Oct. 2023), https://portal.tcu.gov.br/en_us/imprensa/news/tcu-invests-in-institutional-partnerships-data-analysis-and-the-use-of-ai-in-the-fight-against-corruption.htm (accessed 8 Jul. 2024).

³⁰ One notable limitation described by Leonardo Souza-McMurtrie is 'text limitation', meaning the inaccuracy of the AI tool's responses. The author suggests the use of these tools in different languages can also have a drastic impact in their performance. See Leonardo Souza-McMurtrie, *Arbitration Tech Toolbox: Will ChatGPT Change International Arbitration as We Know It?*, Kluwer Arbitration Blog (26 Feb. 2023), <https://arbitrationblog.kluwerarbitration.com/2023/02/26/arbitration-tech-toolbox-will-chatgpt-change-international-arbitration-as-we-know-it> (accessed 9 Jul. 2024).

³¹ As Philip Resnik mentions, 'LLMs trained on purely distributional principles have no way to distinguish among distributional patterns that arise from definitions or meaning, versus normatively acceptable statistical generalisations, versus normatively unacceptable statistical generalisations'. Philip Resnik, *Large Language Models are Biased Because They Are Large Language Models*, arXiv preprint arXiv:2406.13138 (2024), 10.48550/arXiv.2406.13138 (accessed 24 Jun. 2024).

Lack of explainability: Explainability refers to the ability to trace how an AI system uses data to reach a conclusion. Modern GenAI systems, being complex neural networks, often make this impossible, leading to ‘black box’ problems. In decision-making, especially in legal systems, explainability is crucial as it is essential to present the basis for decisions.³²

Hallucinations: Hallucinations in AI refer to producing answers and conclusions that appear plausible but are false or fabricated. GenAI systems, in particular, can provide consistent-seeming yet incorrect responses. Varun Magesh and others describe hallucinations as ‘the tendency of AI tools to produce outputs that are demonstrably false’.³³

3.2 Limitations in Using AI for Decision-Making

In addition to general AI limitations, there are challenges in domain-specific AI systems, such as those used in decision-making.³⁴ Specialized AI systems are being developed, driven by the rise of public GenAI tools. These systems use techniques like fine-tuning and Retrieval Augmented Generation (RAG) with purpose-specific datasets.

Fine-tuning involves taking an existing model that has already been trained on a large, general dataset and then training it further on a more specific, often smaller, dataset related to the particular task at hand. RAG is a technique where the model first retrieves relevant information from a large corpus or database and then uses this information to generate a response. This approach leverages the advantages of both retrieval-based systems, which can provide accurate and factual information, and generation-based systems, which can produce fluent and coherent text. Both techniques use specific datasets, as opposed to the general datasets used to train general LLM models. The more specific the dataset for a certain

matter, the more precise the outcome of the AI system will be.

Decision-making involves at least three types of systems, each with distinct limitations and applications based on their use and primary purpose. First, some systems utilize machine learning and deep learning technology to assess documents and information and to predict outcomes by analysing datasets composed of past cases and precedents. These systems usually offer users one or a few features and utilities. Developed for specific tasks and purposes, they lack flexibility and do not support alternative uses. Similarly, there is limited scope for developing new capabilities or utilities using these systems as a foundation. The first systems developed for e-discovery tasks are prime examples of this type.

Second, there are systems based on GenAI technologies and LLMs. These systems use datasets to generate texts, images, graphics, and sounds, aiding in summarizing, creating text, and translating documents. They employ Generative Pretrained Transformers (GPTs), which predict the desired outcome, whether text, images, sounds, or videos. Their primary function is to produce results according to user prompts. ChatGPT (in its many versions) is the primary example of this kind of system. Other general LLM systems include Gemini, developed by Google, and Claude, developed by Anthropic and the Microsoft copilot system (which is based on the ChatGPT model).

Third, some systems combine the characteristics of both predictive and generative models. These systems can generate content and predict results based on a given dataset. Typically, they are closed and restricted, developed for specific tasks, and designed to serve particular professional and technical niches. Some examples of these systems are Harvey AI,³⁵ CoCounsel,³⁶ Jus Mundi,³⁷ and Lexis+AI.³⁸

There are several limitations to using AI in decision-making contexts.

Notes

³² See Sarah Elisabeth Chojecki and Sophie Nappert: ‘Another problematic aspect of machine learning algorithms is their lack of transparency regarding how and why they can reach a particular output. This phenomenon is commonly referred to as the “black box” feature of AI. To use a simple example, we cannot tell how a self-driving car can navigate through traffic so effectively; however, based on observation, we know it does. As previously noted, in contrast to expert models that follow a deductive pattern of pre-established rules applied to a specific database, machine learning models use pattern recognition to extract algorithms. These models may use hidden factors that are not readily observable, especially when they employ more advanced technologies such as deep learning and neural networks’. Sarah Elisabeth Chojecki & Sophie Nappert, *Evidence in International Arbitration through the Looking-Glass of the Digital Economy*, 8(1) BCDR Int’l Arb. Rev. 121–144 (2021), www.kluwerarbitration.com/document/kli-ka-biar-0801009?q=artificial%20intelligence%20act (accessed 25 Jun. 2024).

³³ V. Magesh, F. Surani, M. Dahl, M. Suzgun, C. D. Manning & D. E. Ho, *Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools* 24 (30 May 2024), <https://arxiv.org/pdf/2405.20362> (accessed 24 Jun. 2024). According to the paper, ‘The most important implication of our results is the need for rigorous, transparent benchmarking and public evaluations of AI tools in law. In other AI domains, benchmarks such as the Massive Multitask Language Understanding (Hendrycks et al., 2020) and BIG Bench Hard (BIG-bench Authors, 2023; Suzgun et al., 2023) have been central to developing a common understanding of progress and limitations in the field. But in contrast to even GPT-4 – not to mention open systems like Llama and Mistral – legal AI tools provide no systematic access, publish few details about models, and report no benchmarking results at all. This stands in marked contrast to the general AI field (Liang et al. 2023), and makes responsible integration, supervision, and oversight acutely difficult. Until progress on these fronts is made, claims of hallucination-free legal AI systems are, at best, ungrounded’.

³⁴ Two of the authors made some prior considerations on this topic. See Cardoso, Pereira, Santini & Souza-McMurtrie, *supra* n. 1.

³⁵ Harvey, www.harvey.ai (accessed 23 Jul. 2024).

³⁶ Co-Counsel, <https://casetext.com> (accessed 23 Jul. 2024). In Aug. 2023, Thomson Reuters acquired Casetext.

³⁷ Jus Mundi, <https://jusmundi.com/en> (accessed 23 Jul. 2024).

³⁸ Lexis+AI, https://law.lexisnexis.com/stanford-digital-ad?utm_source=Google&utm_medium=Search&utm_campaign=SL_2024_Q3_LexisNexis_Stanford_Study_Digital_Ad&utm_term=Stanford&utm_source=google&utm_medium=ppc&utm_term=&utm_campaign=small_law_competitor&gad_source=1&gclid=CjwKCAjw4_K0BhBsEiwAVVZ_4G4lgbC-SjMRfNlfj5tc_MhXx4YPv3hnHYw2sr0WRFEDhGyzQJR0C64QQAvD_BwE (accessed 23 Jul. 2024).

Lack of data: In arbitration, the confidential nature of international commercial arbitration awards restricts the availability of extensive datasets needed to develop specific AI models. Typically, systems are trained on general datasets and fine-tuned using techniques like RAG, incorporating specific datasets that include international norms, treaties, local legislation, awards, and past case data. This limitation affects their use and usefulness.

Biases: A significant concern in decision-making processes is the potential for human biases in the underlying data or methodology used to train AI models. These biases can be perpetuated and amplified in future predictions, increasing the risk of biased outcomes.

Balance between prediction and legal interpretation: Another critical issue is the balance between prediction and legal interpretation. Some AI systems focus on predicting outcomes based on statistics and probabilities rather than providing reasoned legal decisions or understanding the rationale behind past decisions. This could lead to prioritizing outcome prediction over legal reasoning and comprehending the dispute. It may also hinder legal evolution in future decisions.

This challenge is particularly relevant when comparing legal systems. Common law systems rely heavily on precedents, whereas civil law systems place greater importance on legal texts and their interpretation. The distinction between these systems influences the role of AI in the legal field.

AI systems prioritizing outcome prediction may neglect the deeper legal reasoning and rationale behind decisions. This approach can lead to a superficial understanding and application of the law, potentially stagnating legal evolution.

In contrast, AI systems that emphasize legal interpretation can contribute to a richer, more informed legal discourse. In common law, this could mean identifying trends and shifts in judicial reasoning. In civil law, it could enhance the precision and depth of statutory interpretation, promoting dynamic and informed legal development.

Balancing prediction and interpretation is crucial to ensuring that AI systems contribute meaningfully to the legal landscape in both common law and civil law systems.

3.3 Ethical Limitations in Using AI for Decision-Making

It is also important to consider ethical issues when using AI in decision-making.³⁹ Decision-making involves not only legal rules but also ethical standards. Automating this process could complicate the application and enforcement of these standards. The reason is straightforward: ethical rules are difficult to translate into code and systems. Further, ethical principles are often not codified or clearly defined in norms.⁴⁰

National and international entities agree that ethical principles are essential for trustworthy AI systems. The OECD (Organization for Economic Co-operation and Development) has published recommendations on AI, urging members to promote and implement 'principles for responsible stewardship of trustworthy AI'.⁴¹ These principles include:

- (1) Inclusive growth, sustainable development, and well-being
- (2) Respect for the rule of law, human rights, and democratic values, including fairness and privacy
- (3) Transparency and explainability
- (4) Robustness, security, and safety
- (5) Accountability

In 2019, the European Union also published ethical guidelines for trustworthy AI.⁴² The main ethical concerns outlined are:

- (1) Respect for human autonomy
- (2) Prevention of harm
- (3) Fairness
- (4) Explicability

These ethical imperatives could also limit and guide the use of AI tools in decision-making. Many of the principles mentioned, such as fairness, privacy, transparency, accountability, and explainability, are intrinsic to decision-making activities. Using AI tools in decision-making, particularly in dispute resolution, should be in accordance with ethical principles. Both legal rules and ethical principles set boundaries for the employment of AI in these processes.

Notes

³⁹ Leonardo Souza-McMurtrie describes possible ethical risks of AI tools for international arbitration. These include issues of bias, manipulation of the arbitral tribunal, unethical tactics (e.g., misleading or false statements), influence of a ghost writer, and stylistic imitation. See Souza-McMurtrie, *supra* n. 30.

⁴⁰ In this sense, Thomas M. Powers & Jean-Gabriel Ganascia, *The Ethics of the Ethics of AI*, in *The Oxford Handbook of Ethics of AI* (Markus D. Dubber, Frank Pasquale, and Sunit Das eds, Oxford Academic 2020), doi: 10.1093/oxfordhb/9780190067397.013.2 (accessed 5 Jul. 2024). According to the authors, 'the attempts to model ethical reasoning have shown the huge difficulties researchers face in doing so. The first difficulty comes from modelling deontic reasoning, that is, reasoning about obligations and permissions. The second is due to the conflicts of norms that occur constantly in ethical reasoning. The third is related to the entanglement of reasoning and acting, which requires that we study the morality of the act, per se, but also the values of all its consequences' (at 40).

⁴¹ OECD, *Recommendation of the Council on Artificial Intelligence*, OECD/LEGAL/0449 (21 May 2019, Last amended on 2 May 2024), <https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449> (accessed 23 Jun. 2024).

⁴² EU, *EU Ethical Guidelines for Trustworthy AI* (2019), https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419 (accessed 23 Jun. 2024).

Considering the ethical limitations to using AI in decision-making, it is essential that this use considers the core purpose of the activity: decision-making processes are designed and created to solve disputes and ensure social harmony. The use of AI must not compromise this purpose. AI tools should be ethically adopted, ensuring that parties involved in a decision-making process, whether judicial or arbitral, are treated equally and fairly.

All the principles identified above by the OECD and the European Union also reflect ethical principles specific to decision-making. Fairness, accountability, transparency, explainability, and respect for human rights and autonomy can be translated into ethical obligations to be complied with in decision-making processes.

The ethical adoption of AI in decision-making requires that these systems be designed to uphold fairness and equity. This ensures that decisions are not only legally sound but also just and considerate of all parties' rights and circumstances. By adhering to these ethical principles, AI can contribute to more balanced and humane decision-making processes.

4 DISCLOSURE OBLIGATIONS FOR DECISION-MAKERS

4.1 AI and Transparency in the Courts

Courts in various countries are using or experimenting with AI tools and systems in judicial processes. There is a

trend towards adopting AI to enhance the capabilities of judges and judicial officers. This development raises questions about transparency and the need to disclose any use of AI tools. In this section we address how different jurisdictions and international organizations have addressed this issue.

Brazil: Brazil's National Council of Justice (CNJ – Conselho Nacional de Justiça), which oversees the judicial system, has ruled that judges are not prohibited from using AI tools like ChatGPT to assist in decision-making. This decision led the CNJ to reject a motion against the use of AI tools by judges.⁴³ The ruling emphasized that:

the functioning and decision-making criteria of artificial intelligence systems must be transparent and understandable to legal practitioners and the parties involved in judicial processes. Therefore, it is essential to adopt measures to mitigate and monitor algorithmic bias, ensuring fairness and impartiality in judicial decisions.

Article 8 of CNJ Resolution No. 332/2020 establishes specific provisions for transparency standards to be adhered to in the use of AI tools in judicial processes.⁴⁴

The European Union: The Regulation of the European Parliament and of the Council (EU AI Act) laying down harmonized rules on artificial intelligence is the primary legal framework regulating the use of AI tools. Its provisions are based on risk levels and include various mechanisms to ensure the safe use of AI. For this article, it is important to note that the activities of arbitrators, particularly decision-making, may be classified as 'high-risk' under the AI Act.⁴⁵

Notes

⁴³ Decision in the process no. 0000416–89.2023.2.00.0000, rendered on 25 Jun. 2024, www.conjur.com.br/wp-content/uploads/2024/07/9769C290CD7A61_0000416-89.2023.2.00.0000_5617.pdf (accessed 27 Jun. 2024). According to the rapporteur's opinion:

In the context of the Brazilian Judiciary, the use of AI is addressed in CNJ Resolution No. 332/2020, which covers ethics, transparency, and governance in the development and use of artificial intelligence. This Council's regulation acknowledges that AI should be used to 'promote and deepen a greater understanding between the law and human actions, between freedom and judicial institutions' (Art. 1). The inventiveness of this technology drives a significant step forward in the evolution of generative artificial intelligence.

Nonetheless, the use of tools like ChatGPT and other AI technologies by the Judiciary requires careful analysis in light of various ethical, legal, and constitutional principles. These tools offer significant potential to enhance the efficiency and effectiveness of the judicial system, but their application demands specific attention to ethics, fairness, and responsibility in their use.

⁴⁴ According to Art. 8 of CNJ Resolution No. 332/2020:

For the purposes of this Resolution, transparency consists of: I – Responsible disclosure, considering the inherent sensitivity of judicial data; II – Indication of the objectives and intended outcomes of using the Artificial Intelligence model; III – Documentation of identified risks and indication of the information security and control instruments to address them; IV – Possibility of identifying the reason in case of damage caused by the Artificial Intelligence tool; V – Presentation of audit and certification mechanisms for best practices; VI – Provision of a satisfactory and auditable explanation by a human authority for any proposed decision presented by the Artificial Intelligence model, especially when it is of a judicial nature.

⁴⁵ According to recital (61) of the AI Act,

Certain AI systems intended for the administration of justice and democratic processes should be classified as high-risk, considering their potentially significant impact on democracy, the rule of law, individual freedoms as well as the right to an effective remedy and to a fair trial. In particular, to address the risks of potential biases, errors and opacity, it is appropriate to qualify as high-risk AI systems intended to be used by a judicial authority or on its behalf to assist judicial authorities in researching and interpreting facts and the law and in applying the law to a concrete set of facts.

AI systems intended to be used by alternative dispute resolution bodies for those purposes should also be considered to be high-risk when the outcomes of the alternative dispute resolution proceedings produce legal effects for the parties. The use of AI tools can support the decision-making power of judges or judicial independence, but should not replace it: the final decision-making must remain a human-driven activity.

According to Annex III, 8, (a) of the AI Act, ‘AI systems intended to be used by a judicial authority or on their behalf to assist a judicial authority in researching and interpreting facts and the law and in applying the law to a concrete set of facts, or to be used in a similar way in alternative dispute resolution’ are classified as high-risk AI systems and shall comply with the related provisions related.

New Zealand: The New Zealand Guidelines for the Use of Generative AI in Courts and Tribunals provide a framework for judges, judicial officers, tribunal members, and judicial support staff.⁴⁶ These guidelines emphasize the importance of understanding the capabilities and limitations of generative AI chatbots, such as ChatGPT, Bing Chat, and Google Bard.⁴⁷ Users are cautioned about generative AI’s potential inaccuracies, biases, and limited access to New Zealand law and procedural requirements.

Under the guidelines, confidentiality, privacy, and suppression are paramount; hence, sensitive information should not be entered into generative AI chatbots due to the risk of unintended public disclosure. Users must verify the accuracy of GenAI-generated information before relying on it, be mindful of ethical considerations, including biases in training data, and adhere to copy-right laws.

Notably, judicial officers need not disclose generative AI usage routinely.⁴⁸ However, clerks, research counsel, and judicial support staff are advised to discuss using generative AI tools with their supervising judicial officers to mitigate risks.

United Kingdom: The UK Guidance for Judicial Office Holders on Using AI stresses responsible AI usage while highlighting disclosure obligations.⁴⁹ It advises against entering any private or confidential information into public AI chatbots, cautioning that

such data could be inadvertently made public.⁵⁰ It recommends disabling chat history where possible and being cautious about device app permissions to mitigate risks.

It also outlines steps for reporting unintentional disclosure of sensitive information, emphasizing the judiciary’s duty to safeguard the integrity of justice administration. This guidance underscores judicial officeholders’ need to understand AI tools’ capabilities and limitations, ensuring any AI-assisted material is verified for accuracy and appropriateness before use.⁵¹

Similar to the New Zealand guidelines, the UK guidelines state that judges are not generally obliged to describe the research or preparatory work which may have been done in order to produce a judgment.⁵² The guidelines state that, provided the guidelines are appropriately followed, there is no reason why generative AI could not be a potentially useful secondary tool. If court clerks, judicial assistants, or other team members use AI technologies during their duties, judges should engage in dialogue with them about such use.

The United Nations: In March 2024, the United Nations issued a resolution titled ‘Seizing the Opportunities of Safe, Secure, and Trustworthy Artificial Intelligence Systems for Sustainable Development’ (Resolution).⁵³ The UN encourages Member States to promote safe, secure, trustworthy, inclusive, and equitable AI systems that benefit all.⁵⁴

The Resolution advocates for transparency, predictability, reliability, and understandability throughout the AI system lifecycle, especially in decision-making processes. It emphasizes the importance of providing notice and explanation, promoting human oversight, reviewing auto-

Notes

⁴⁶ Courts of New Zealand, ‘New Zealand Guidelines for Use of Generative AI in Courts and Tribunals: Judges, Judicial Officers, Tribunal Members and Judicial Support Staff’ (7 Dec. 2023), www.courtsnz.govt.nz/going-to-court/practice-directions/practice-guidelines/all-benches/guidelines-for-use-of-generative-artificial-intelligence-in-courts-and-tribunals (accessed 8 Jul. 2024).

⁴⁷ *Ibid.*, at 2.

⁴⁸ *Ibid.*, at 3.

⁴⁹ UK Courts and Tribunals Judiciary, *Artificial Intelligence (AI): Guidance for Judicial Office Holders 2* (12 Dec. 2023), www.judiciary.uk/wp-content/uploads/2023/12/AI-Judicial-Guidance.pdf (accessed 19 Jul. 2024).

⁵⁰ *Ibid.*, at 3.

⁵¹ *Ibid.*, at 2 and 5.

⁵² *Ibid.*, at 5.

⁵³ United Nations, Resolution on Seizing the opportunities of safe, secure, and trustworthy artificial intelligence systems for sustainable development. Document A/78/L.49 (11 Mar. 2024) (UN Resolution A/78/L.49), <https://daccess-ods.un.org/access.nsf/Get?OpenAgent&DS=A/78/L.49&Lang=E> (accessed 25 Jun. 2024).

⁵⁴ Article 6 of UN Resolution A/78/L.49 states

(a) Promoting the development and implementation of domestic regulatory and governance approaches and frameworks, in line with their respective national, and where applicable subnational, policies and priorities and obligations under international law, to support responsible and inclusive artificial intelligence innovation and investment for sustainable development, while simultaneously promoting safe, secure and trustworthy artificial intelligence systems.

mated decisions, and ensuring human decision-making alternatives or effective redress for those adversely impacted by AI decisions.⁵⁵

The Resolution also mandates that AI systems must comply with human rights. It states that 'human rights and fundamental freedoms must be respected, protected, and promoted throughout the lifecycle of artificial intelligence systems'. It calls on Member States and stakeholders to refrain from using AI systems that cannot operate in compliance with international human rights law or pose undue risks to human rights, particularly for those in vulnerable situations. The Resolution reaffirms that people's offline rights must be protected online, including throughout the AI system lifecycle (Article 5).⁵⁶

Given that due process can be considered a human right, AI tools and systems must comply with due process obligations. This includes adhering to transparency requirements to protect these rights.

The United States: The United States adopts a similar approach. Although there is no specific legal provision regulating AI or establishing limits on AI systems, the federal government has initiated measures related to AI technology.

In 2022, the U.S. Federal Government issued the Blueprint for an AI Bill of Rights.⁵⁷ The AI Bill of Rights outlines key principles essential for the ethical use of AI in legal decision-making. It emphasizes the necessity for safe and reliable AI systems to prevent harm. To avoid algorithmic discrimination, AI

must be designed to prevent reinforcing biases and ensure fair outcomes. Data privacy is paramount, granting individuals control over their personal information and protection against misuse. Transparency is crucial, requiring clear explanations and notifications about AI decisions. Additionally, human alternatives and fallback options must be available to maintain trust and accountability in AI-driven processes.⁵⁸

In October 2023, the Federal Government issued Executive Order No. 14110 on Safe, Secure, and Trustworthy Artificial Intelligence.⁵⁹ The Executive Order on Safe, Secure, and Trustworthy AI establishes key principles relevant to legal decision-making. It emphasizes the need to prevent algorithmic discrimination, ensuring AI systems do not reinforce biases. It mandates transparency and accountability, requiring clear explanations for AI-driven decisions. The order prioritizes privacy protection and safeguarding personal data to maintain confidentiality in legal contexts. Additionally, it guides the use of AI to promote equity in criminal justice and ensures that human oversight and alternatives are available in AI-driven legal processes.⁶⁰

In 2024, the Office of Management and Budget (OMB) published memorandum M-24-10, directing agencies to 'advance AI governance and innovation while managing risks from the use of AI in the federal government, particularly those affecting the rights and safety of the public'.⁶¹

Notes

⁵⁵ Article 6 of UN Resolution A/78/L.49 states:

(k) Promoting transparency, predictability, reliability and understandability throughout the life cycle of artificial intelligence systems that make or support decisions impacting end-users, including providing notice and explanation, and promoting human oversight, such as, for example, through review of automated decisions and related processes or, where appropriate and relevant, human decision – making alternatives or effective redress and accountability for those adversely impacted by automated decisions of artificial intelligence systems.

⁵⁶ Article 5 of UN Resolution A/78/L.49 states:

Emphasizes that human rights and fundamental freedoms must be respected, protected and promoted throughout the life cycle of artificial intelligence systems, calls upon all Member States and, where applicable, other stakeholders to refrain from or cease the use of artificial intelligence systems that are impossible to operate in compliance with international human rights law or that pose undue risks to the enjoyment of human rights, especially of those who are in vulnerable situations, and reaffirms that the same rights that people have offline must also be protected online, including throughout the life cycle of artificial intelligence systems.

⁵⁷ U.S. Federal Government, *The Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People* (2022), www.whitehouse.gov/ostp/ai-bill-of-rights (accessed 24 Jun. 2024).

⁵⁸ According to the Blueprint for an AI Bill of Rights, 'In some cases, a human or other alternative may be required by law. You should have access to timely human consideration and remedy by a fallback and escalation process if an automated system fails, it produces an error, or you would like to appeal or contest its impacts on you. Human consideration and fallback should be accessible, equitable, effective, maintained, accompanied by appropriate operator training, and should not impose an unreasonable burden on the public. Automated systems with an intended use within sensitive domains, including, but not limited to, criminal justice, employment, education, and health, should additionally be tailored to the purpose, provide meaningful access for oversight, include training for any people interacting with the system, and incorporate human consideration for adverse or high-risk decisions'.

⁵⁹ U.S. Federal Government, 'Fact Sheet: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence' (30 Oct. 2023), www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/ (accessed 23 Jun. 2024).

⁶⁰ See specially these items of the Executive Order: 'Address algorithmic discrimination through training, technical assistance, and coordination between the Department of Justice and Federal civil rights offices on best practices for investigating and prosecuting civil rights violations related to AI. Ensure fairness throughout the criminal justice system by developing best practices on the use of AI in sentencing, parole and probation, pretrial release and detention, risk assessments, surveillance, crime forecasting and predictive policing, and forensic analysis' (emphasis added).

⁶¹ US Federal Government, *Memorandum for the Heads of Executive Departments and Agencies* (28 Mar. 2024), www.whitehouse.gov/wp-content/uploads/2024/03/M-24-10-Advancing-Governance-Innovation-and-Risk-Management-for-Agency-Use-of-Artificial-Intelligence.pdf (accessed 22 Jun. 2024).

The Memorandum on Advancing Governance, Innovation, and Risk Management for Agency Use of AI contains relevant provisions for decision-making processes. It introduces essential guidelines to enhance legal decision-making, establishing comprehensive AI governance frameworks to ensure responsible and ethical deployment in legal contexts. By implementing risk management strategies it aims to mitigate biases and potential harms, ensuring fairness and justice in AI-driven decisions.⁶²

The memorandum also emphasizes the importance of transparency in AI operations, fostering trust and accountability. Strengthened accountability mechanisms ensure rigorous oversight, while adherence to ethical standards guarantees the integrity of legal processes.⁶³ A relevant provision pertains to the obligation to provide explanations for an agency's decisions and actions. The Memorandum states that:

Explanations might include, for example, how and why the AI-driven decision or action was taken. This does not mean that agencies must provide a perfect breakdown of how a machine learning system came to a conclusion, as exact explanations of AI decisions may not be technically feasible.

However, agencies should still characterize the general nature of such AI decisions through context such as the data that the decision relied upon, the design of the AI, and the broader decision-making context in which the system operates.

Such explanations should be technologically valid, meaningful, useful, and as simply stated as possible, and higher-risk decisions should be accompanied by more comprehensive explanations.⁶⁴

4.2 AI and Transparency in Arbitration

In the evolving landscape of arbitration and litigation, the balance between protecting sensitive information

and ensuring transparency is becoming increasingly important. As technology, particularly artificial intelligence, becomes more integrated into legal processes, various guidelines have been established to navigate these complexities.⁶⁵ This section discusses the approaches taken by different organizations and institutions to address the use of AI in legal contexts, emphasizing the need for disclosure, ethical considerations, and maintaining the integrity of legal proceedings.

SVAMC Guidelines⁶⁶: These Guidelines, opened for public consultation in August 2023 and published on 30 April 2024, do not impose mandatory disclosure by default.⁶⁷ Rather, they provide that 'decisions regarding disclosure of the use of AI tools shall be made on a case-by-case basis taking account of the relevant circumstances, including due process and any applicable privilege'.⁶⁸

The SVAMC Guidelines acknowledge that some AI uses may be uncontroversial, not necessitating disclosure. However, they open the door for arbitral tribunals, parties, or administering institutions to require disclosure and/or certification related to the use of Generative AI tools, depending on the circumstances. These circumstances may include evolving technology accuracy and the specific AI tool employed. This need for disclosure is contingent on circumstances that could affect the integrity of proceedings or evidence.

(American Arbitration Association – International Centre for Dispute Resolution) Principles for AI in ADR (AAA-ICDR Principles): The AAA-ICDR Principles for AI in ADR are another set of guidelines for arbitration published in November 2023.⁶⁹ They aim to integrate AI into ADR, enhancing efficiency while upholding legal integrity and service. These principles underscore the importance of maintaining competence, confidentiality, advocacy, impartiality, independence, and process improvement when using AI.

Notes

⁶² See item v.A.4 'Consistent with applicable law, cease use of the AI for agency decision-making if the agency is unable to adequately mitigate any associated risk of unlawful discrimination against protected classes. Agencies should maintain appropriate documentation to accompany this decision-making, and should disclose it publicly to the extent consistent with applicable law and governmentwide policy' (at 21).

⁶³ See e.g., provision H: 'Provide additional human oversight, intervention, and accountability as part of decisions or actions that could result in a significant impact on rights or safety. Agencies must assess their rights-impacting and safety-impacting uses of AI to identify any decisions or actions in which the AI is not permitted to act without additional human oversight, intervention, and accountability. When immediate human intervention is not practicable for such an action or decision, agencies must ensure that the AI functionality has an appropriate fail-safe that minimizes the risk of significant harm' (at 20).

⁶⁴ *Ibid.*, at 23, footnote n. 50.

⁶⁵ Two of the authors previously analysed the relevant guidelines relating to AI in legal decision-making. See Cardoso, Pereira, Santini & Souza-McMurtrie, 'supra' n. 1.

⁶⁶ One of the authors, Elizabeth Chan, is on the drafting committee of SVAMC Guidelines.

⁶⁷ Silicon Valley Arbitration and Mediation Center Guidelines on the Use of Artificial Intelligence in Arbitration (30 Apr. 2024), Guideline 3, <https://svamc.org/wp-content/uploads/SVAMC-AI-Guidelines-First-Edition.pdf> (accessed 8 Jul. 2024).

⁶⁸ *Ibid.*

⁶⁹ AAA & ICDR, *Principles Supporting the Use of AI in Alternative Dispute Resolution* (Nov. 2023), <https://mediate.com/aaa-icdr-principles-supporting-the-use-of-ai-in-alternative-dispute-resolution/> (accessed 8 Jul. 2024).

They emphasize the critical need for legal professionals to be proficient in AI technologies, understand associated risks and benefits, and adhere to ethical considerations, ensuring that AI integration does not compromise confidentiality or the quality of arbitration and mediation.⁷⁰ These Principles do not expressly address the issue of disclosure.

MIT Guidelines: The Massachusetts Institute of Technology (MIT) Task Force on the Responsible Use of AI for Law focuses on maintaining high ethical and professional standards in using generative AI within legal contexts generally.⁷¹

The guidelines underscore seven key duties: confidentiality, fiduciary care, client notice and consent, competence, fiduciary loyalty, regulatory compliance, and accountability and supervision.⁷²

These principles are designed to ensure that the deployment of AI tools in legal work respects the confidentiality of client information, upholds fiduciary responsibilities, and complies with regulatory norms while emphasizing the necessity for human oversight and accountability in the use of AI applications.⁷³

Central to these guidelines is the implicit need for disclosure, highlighted by the duty of accountability and (human) supervision.⁷⁴ This duty necessitates maintaining human control over AI usage and outputs to ensure the integrity of legal processes and the accuracy of information provided to clients and courts.⁷⁵

While explicit mandates on disclosure are not detailed in the principles, the emphasis on accountability and supervision suggests a framework where legal professionals may need to disclose their use of AI tools under certain conditions to uphold transparency and trust in legal proceedings.⁷⁶ The MIT task force advocates for a balanced use of AI in law, recognizing the potential of AI to enhance legal practice while ensuring that its integration does not compromise professional obligations or the welfare of clients.⁷⁷

JAMS (formerly Judicial Arbitration and Mediation Services, Inc.) AI Rules: The JAMS AI Rules introduce a novel procedure for producing and inspecting AI systems and related materials to safeguard sensitive and

confidential information, including trade secrets inherent in AI technologies. Rule 16.1(b) stipulates that the disclosing party must make these materials – encompassing hardware, software, models, and training data – available to one or more experts within a secured environment. These experts are prohibited from transmitting or removing any materials from this environment.

Practically, this means that the requesting party and its counsel do not directly access the disclosed information but must rely on the expert's review conducted within the secure setting established by the disclosing party.⁷⁸ This procedure aims to ensure the protection of sensitive AI data while allowing necessary expert analysis.

5 CHALLENGES IN AI-ASSISTED DECISION-MAKING

As AI becomes more integrated into arbitration, this section explores the complexities and difficulties of AI-assisted decision-making. It examines the possible challenges involving the use of AI in arbitration, emphasizing the need for transparency and clear boundaries. It also discusses the importance of arbitrators being technically competent to handle AI tools effectively and highlights strategies to address these challenges, including education, updates to legal frameworks, and the development of guidelines to ensure fairness, transparency, and accountability.

5.1 Overcoming Possible Challenges of AI-Assisted Decision-Making in Arbitration

Understanding AI-assisted decision-making involves scrutinizing both the procedure and the outcome of the decision-making process. Focusing on arbitration, a fundamental characteristic of it is the parties' agreement to submit their dispute to an impartial third party. At the heart of the arbitration process is the parties' choice to have a private decision-maker resolve their dispute according to specific rules and procedures. The selection of arbitrators is also a direct or indirect result of the parties' will.

Notes

⁷⁰ *Ibid.*, at 1.

⁷¹ Dazza Greenwood, *Task Force on Responsible Use of Generative AI for Law*, MIT Computational Law Report 5 (28 Feb. 2023), <https://law.mit.edu/pub/generative-ai-responsible-use-for-law> (accessed 8 Jul. 2024).

⁷² *Ibid.*, at 5.

⁷³ *Ibid.*, at 6.

⁷⁴ *Ibid.*

⁷⁵ *Ibid.*, at 3.

⁷⁶ *Ibid.*, at 3.

⁷⁷ *Ibid.*, at 2.

⁷⁸ Sam Winter-Barker, *JAMS Publishes Artificial Intelligence Arbitration Rules, but Are They Fit for Purpose?*, Kluwer Arbitration Blog (27 Jun. 2024), <https://arbitrationblog.kluwerarbitration.com/2024/06/27/jams-publishes-artificial-intelligence-arbitration-rules-but-are-they-fit-for-purpose> (accessed 14 Jul. 2024).

When parties include an arbitration clause in an agreement, they express a clear preference for defining, in detail, how disputes arising from the agreement will be resolved. This choice underscores the voluntary nature of arbitration, distinguishing it from state courts. Unlike litigation, governed by strict legal procedures, arbitration allows parties significant control over managing and resolving their disputes. This flexibility and autonomy are central to the appeal of arbitration. Even in exceptional cases where arbitration is mandatory by statute, parties still retain substantial control over the process.

This premise is central to the possible challenges arising from AI-assisted decision-making in arbitration. When parties voluntarily choose arbitration, they seek certainty and control over aspects of the decision-making process. The flexibility in defining procedural aspects of arbitration underscores this intent.

As a rule, the parties involved should have some say on whether and how AI will be employed in arbitration processes. They may agree on the use of AI by the arbitration tribunal, or at the very least, be aware of its use and satisfied that it will comply with what is expected as a fair process. This approach addresses the issue directly. The most evident ground for challenging an arbitration procedure or award is when the arbitral tribunal fails to comply with procedural fairness in the use of AI-assisted decision-making tools during the arbitration.

Transparency about whether and how to employ such tools is a useful (and arguably necessary) step to avoid possible challenges. Arbitrators must be mindful that the lack of specialized knowledge on the part of the parties and counsel may also play a role. It may lead parties and counsel to be insecure regarding the use of AI tools by the arbitral tribunal and increase the risk of challenges, either founded or not. Openness and transparency are key.

Moreover, there are nuances to this discussion. Even if the parties agree to the use of AI-assisted tools in arbitration, challenges can still arise. These challenges often stem from a lack of clarity regarding the limits and extent of AI usage. For example, parties might agree to use AI tools to expedite the process by summarizing arguments or transcribing oral hearings. However, they may not consent to AI involvement in the actual decision-making process.

It is in these grey areas that significant difficulties in the use of AI may emerge. Therefore, it is crucial for parties to clearly understand and explicitly define the specific roles and boundaries of AI in the specific arbitration process to avoid potential disputes or misunderstandings. The widespread lack of specialized knowledge among arbitrators, parties and counsel only adds to the complexity of the situation. An open discussion about the use of AI among the arbitral tribunal, parties and counsel

at the outset of the arbitration (or whenever the issue of possible AI use comes up) is the best path to avoid difficulties down the road.

Potential challenges to AI-assisted decisions may mirror concerns over tribunal secretaries' involvement. AI tools typically play a role similar to that of tribunal secretaries. In any scenario, the core issue revolves around transparency, accountability, and the extent of influence on the decision-making process.⁷⁹

First, transparency and disclosure are crucial in addressing concerns about AI in arbitration.⁸⁰ Similar to worries about undisclosed tribunal secretaries, parties are apprehensive about the use of AI. They should be fully informed about how extensively AI tools are employed for administrative tasks, procedural support, or substantive decision-making. Without transparency, mistrust can arise, leading to challenges to the arbitration's validity.

Second, the extent of influence is another critical factor to consider. Tribunal secretaries are traditionally assumed to provide administrative assistance, but concerns arise when their involvement extends to influencing or drafting substantive parts of the award. Similarly, AI tools may be assumed to be acceptable for tasks like data organization or document management. However, if AI is perceived to influence or make substantive decisions without the parties' explicit consent, it could undermine the legitimacy of the arbitration process.

Third, preserving and enhancing accountability is crucial. Tribunal secretaries, though human, act under the direct supervision of arbitrators, who remain accountable for the final award. With AI, accountability becomes more complex. Ensuring that arbitrators remain fully responsible for decisions, even when using AI tools, is essential to maintaining confidence in the process. Arbitrators must be able to explain and justify their decisions, regardless of AI assistance.

Fourth, information and consent are fundamental. Similar to the involvement of tribunal secretaries, the use of AI tools should be based on the parties' informed consent. Parties must be clearly aware of the scope and limitations of AI involvement. This agreement helps prevent future disputes and challenges based on alleged overreach or unreasonable use of AI in the decision-making process.

Fifth, addressing perceived bias and fairness is essential. Concerns about bias and fairness are prevalent in both tribunal secretaries and AI contexts. Parties might worry that tribunal secretaries could introduce unintended biases. Similarly, there is apprehension that AI algorithms, if not properly vetted and understood (and supervised), might perpetuate or even exacerbate biases. Ensuring that AI tools are transparent, unbiased, and used appropriately is crucial to maintaining fairness.

Notes

⁷⁹ Cardoso, Pereira, Santini & Souza-McMurtrie, *supra* n. 1.

⁸⁰ Disclosure is required by a number of institutional rules, as discussed in this article.

In summary, the potential challenges to AI-assisted decisions may mirror those related to tribunal secretaries' involvement. Both require clear guidelines, transparency, and informed consent to ensure that the arbitration process remains fair, accountable, and trusted by all parties involved.

Conversely, misconceptions or prejudices should not bar either one from being used, as both AI tools and tribunal secretaries can enhance efficiency in arbitration. Establishing robust protocols and maintaining open communication about the use of AI tools, just as with tribunal secretaries, can help mitigate concerns and uphold the integrity of the arbitration process. Published guidelines and protocols may be an efficient tool to be incorporated into specific arbitral proceedings as an agreed-upon benchmark for the conduct of the parties, counsel and the arbitral tribunal.

Additionally, maintaining adjudicative authority exclusively within the tribunal is paramount to preserving the integrity and legitimacy of the arbitration process. While AI-assisted tools can enhance efficiency and support administrative tasks, the ultimate responsibility for decision-making must reside with the arbitrators. This ensures that the nuanced understanding, human judgment, and accountability that characterize a fair and just arbitration process are preserved. Arbitrators, not AI, must interpret the facts, apply the law, and deliver the final award to guarantee that decisions are rendered impartially and transparently.

By keeping adjudicative authority firmly within the tribunal, parties can be confident that qualified individuals, fully accountable for the outcomes, are resolving their disputes. This approach safeguards the fundamental principles of arbitration. Furthermore, case law regarding tribunal secretaries may help dispel unnecessary concerns about the use of AI in the arbitration process.⁸¹

5.2 Arbitrators' duty to be Technically Competent?

The lack of technical expertise among arbitrators in using AI tools is another difficult challenge to overcome. Even if it is not in itself a ground for challenge, it may lead to the mishandling of AI tools and their use in arbitral proceedings. Arbitrators who responded to the Bryan Cave

Leighton Paisner Annual Arbitration Survey 2023: The Rise of Machine Learning expressed limited confidence in their technical capabilities to adjudicate issues involving AI tools. Specifically, 73% of arbitrator respondents rated their confidence in their technical abilities at five or below on a scale of 1 to 10, with only 3% feeling very confident.⁸² The broader survey respondents echoed this sentiment, with 79% rating their confidence in arbitrators' technical capability to direct the use of AI tools in arbitration at five or below.⁸³

Arbitrators' self-assessed lack of technical knowledge has broader implications for the arbitration process. Their lack of confidence in handling issues related to AI use can lead to challenges in ensuring fairness, transparency, and due process. Sole practitioners are particularly disadvantaged, as they are less likely to have access to AI tools specifically designed for legal practitioners, with appropriate ring-fencing and confidentiality protections, compared to practitioners in large firms who benefit from economies of scale.

In *The Technological Competence of Arbitrators: A Comparative and International Legal Study*, Katia Fach Gomez argues that arbitrators must possess technological competence.⁸⁴ She defines this competence as the ability to perform tasks efficiently and highlights the need for arbitrators to have technological skills in addition to legal expertise. Gomez distinguishes between basic competence, a fundamental expectation for all arbitrators, and premium competence, which offers a competitive edge for handling complex cases.

Although there is no explicit duty for technological competence, broader legal frameworks and ethical guidelines imply its necessity. For example, Gomez references domestic rules of conduct requiring lawyers to stay informed about technological advancements. Gomez suggests extending such obligations to arbitrators, emphasizing the importance of technological acuity in maintaining fairness, efficiency, and expediency in arbitration.

Such an implied duty (or expectation) may not in itself be a ground for challenging an arbitrator. However, it can lead to difficulties in adequately handling the available AI tools, and such mishandling could amount to relevant procedural defects. An arbitrator unfamiliar with the technology may inadvertently jeopardize the legitimacy and validity of the arbitration process and its outcome.

Notes

⁸¹ *European Commission v. Emek and WTE*, Case C.21.0548.F/1, Belgian Court of Cassation (24 Apr. 2023) ('En l'espèce, il ressort des pièces auxquelles votre Cour peut avoir égard que la secrétaire administrative était chargée de rédiger des questions lors de l'audition des experts et a participé à la rédaction de la sentence arbitrale. Il en résulte que tant le tribunal arbitral que la secrétaire administrative ont méconnu les limites qui s'imposent à leur mission dès lors que: – les arbitres n'ont pas respecté le caractère intuitu personae de leur mission; – la secrétaire administrative s'est vue attribuer des fonctions juridictionnelles'). In a comment on the decision, authors André Guskow and Cesar Pereira suggested that 'AI should remain an assistant, under the ultimate discretion of the human arbitrator' (Cardoso, Pereira, Santini & Souza-McMurtrie, *supra* n. 1).

⁸² Bryan Cave Leighton Paisner, Annual Arbitration Survey 2023: The Rise of Machine Learning 20 (9 Nov. 2023), www.bclplaw.com/a/web/tUW2SW6fjHrpXVrA7AfWkS/102932-arbitration-survey-2023-report_v10.pdf (accessed 23 Jul. 2024).

⁸³ *Ibid.*, at 20.

⁸⁴ Katia Fach Gomez, *The Technological Competence of Arbitrators: A Comparative and International Legal Study* (Springer 2023).

5.3 Addressing Challenges in AI-Assisted Decision-Making in Arbitration

Mitigating challenges in AI-assisted decision-making in arbitration requires a multi-faceted approach, which could include, for example:

- (1) Educating arbitrators, legal practitioners, and parties involved in arbitration about AI tools and their implications. Training programs can ensure that all stakeholders understand the technology, its benefits, and potential risks, enabling informed decisions and upholding due process.
- (2) Introducing a 'tech-era due process clause' could be beneficial as arbitration increasingly integrates AI and other technological advances. This evolving concept of due process would aim to ensure that parties' rights are protected in a digital context, maintaining fairness, transparency, and accountability in AI-assisted decision-making.

An evolving concept of due process could become a basis for challenging AI-assisted decisions, such as in award enforcement contexts. Issues could include lack of transparency, absence of informed consent, evidence of AI bias, or failure to maintain human oversight. Implementing procedural safeguards, like independent audits of AI tools and establishing channels for raising concerns, can help address challenges effectively. These safeguards ensure that parties have recourse if AI usage has compromised due process.

- (3) Updating legal frameworks governing arbitration to include explicit provisions related to AI and technology (e.g., in procedural orders or arbitration rules) could also be considered. This could involve drafting new regulations or amending current ones to incorporate transparency, informed consent, accountability, fairness, and data security principles.

Developing guidelines and best practices, such as those from the SVAMC, can help standardize AI integration in arbitration. These guidelines should cover all aspects of AI use, from initial consent to ongoing oversight and dispute resolution mechanisms.

In summary, adapting the due process clause to encompass technological advances in AI involves updating legal frameworks, developing guidelines, providing education, and implementing robust safeguards.

6 CASE STUDY: CHATGPT ARBITRATORS IN THE VIS MOOT

In this section, we explore the practical application of AI in arbitration through a case study using ChatGPT as an arbitrator in the Vis Moot. Subsection 6.1 details the premises of the experiment, including the setup and roles assigned to ChatGPT and human participants. Subsection 6.2 discusses the preparation process, highlighting the creation and configuration of specific ChatGPT accounts for the experiment. Subsection 6.3 provides an overview of the experimental hearing itself, including the interactions and responses of ChatGPT. Finally, subsection 6.4 reflects on the lessons learned from the experiment, focusing on the limitations of AI technology, due process concerns, and potential enforceability issues:

6.1. Premises of the Experiment: ChatGPT as Wing Arbitrators, Humans as TRIBUNAL CHAIR and Party Advocates, 30th Vis Moot Problem as the Case

We aimed to test whether ChatGPT could make decisions in a case in the same way as human arbitrators and to what extent ChatGPT could potentially replace arbitrators.

We conducted live experimental hearings with the support of the Chartered Institute of Arbitrators Brazil Branch (CI Arb Brazil), the Association for the Organization and Promotion of the Willem C. Vis International Commercial Arbitration Moot,⁸⁵ the Vis East Moot Foundation, the Vienna International Arbitration Centre (VIAC), and the Hong Kong law firm Tanner De Witt. These experiments took place in Hong Kong⁸⁶ and Vienna⁸⁷ during the 2024 Vis East and Vis Moot, respectively.

We designed our experimental arbitration hearing to involve a commercial dispute between two parties, resolved by a three-member arbitral tribunal. ChatGPT, prompted by humans, played the roles of the wing arbitrators. A human chaired the tribunal, and human advocates represented the parties.

For the case, we used the 30th Willem C. Vis International Commercial Arbitration Moot (20th Willem C. Vis (East) Moot Problem) (the Problem). The Problem involved a purchase and supply agreement for Unmanned Aerial Systems (UAS) or drones between Drone Eye plc (the Claimant) and Equatoriana Geoscience Ltd (the Respondent), from the fictitious countries of Mediterraneo and Equatoriana, respectively.

Notes

⁸⁵ Verein zur Veranstaltung und Förderung des Willem C. Vis International Commercial Arbitration Moot, an Austrian non-profit association.

⁸⁶ The entire experiment held in Hong Kong is available at, <https://youtu.be/K3Yvz1T4AiI?si=m7HD5xEgYsPNzetA> (accessed 9 Jul. 2024).

⁸⁷ The entire experiment held in Vienna is available at: <https://youtu.be/CrGejdZhpIs?si=h4cclFe5M79VKfUy> (accessed 9 Jul. 2024).

The Respondent alleged corruption and misrepresentation regarding the drones' quality and capabilities, leading them to declare the contract void and terminate negotiations. The agreement contained an arbitration clause invoking the rules of the Permanent Court of Arbitration (PCA), and the Claimant initiated arbitration proceedings.

The main legal questions in the Problem were:

- (1) Does the Arbitral Tribunal have jurisdiction to hear the dispute?
- (2) If the Tribunal's jurisdiction is established, should the proceedings be stayed until the investigations against Mr Field are concluded or bifurcated?
- (3) Is the Purchase and Supply Agreement governed by the United Nations Convention on Contracts for the International Sale of Goods (CISG)?
- (4) If the Purchase and Supply Agreement is governed by the CISG, can the Respondent rely on Article 3.2.5 of the International Commercial Contract Act of Equatoriana to avoid the contract, or is the Claimant correct that this is excluded based on the facts?

Given the Problem involved two procedural questions and two substantive questions, we conducted two experimental hearings: one in Hong Kong (focusing on questions 1 and 3) and another in Vienna (focusing on questions 2 and 4).

For this article, we refer to the experimental hearing conducted in Hong Kong in April 2024, in which three authors (Elizabeth Chan, Luísa Quintão, and Cesar Pereira) were personally involved.

In that hearing, Australian arbitrator Mariel Dimsey served as the chair of the arbitral tribunal; Luísa Quintão and Cesar Pereira prompted the ChatGPT wing arbitrators; Elizabeth Chan gave an introductory speech on AI and dispute resolution; and students Echo Cheng, Janis Winkler, Tino Wasner, and Yongwen Zheng played the roles of the parties' advocates. An active Q&A session with the audience followed. The recording of the Hong Kong experiment is available online.⁸⁸

6.2 Preparation for the Experiment: Setting up ChatGPT Accounts, Creating Specific Arbitrator GPTs, and Testing the Set-up

In preparation for the experimental hearing, we created specific accounts and fine-tuned the GPT-4 model to act

as wing arbitrators, which we named Neo and Trinity. We tailored these GPTs using a specific data set for their roles as arbitrators.

ChatGPT allows users to configure their GPTs by setting their name, description, instructions, and conversation starters. During configuration, users can upload 'knowledge' files and filter the GPTs' capabilities, such as web browsing, image generation, code interpretation, and data analysis, among other settings.

We created a specific GPT named '30th Vis Moot Arbitrator' within the Neo and Trinity accounts. Its general description instructed it to 'act as an arbitrator for an international commercial arbitration based on the Problem of the 30th Willem C. Vis International Commercial Arbitration Moot'.

We also provided the GPT with specific instructions concerning how we expected the model to act:

Activities: Evaluate arguments presented, ask for clarifications, assess the quality of the arguments and their applicability to the case, provide constructive feedback, and decide the correct outcome of the case.

Conduct: You should always refer to the facts of the case, case law, scholarly opinions, and the memoranda you have been provided with and, as applicable, browse the internet for further information. Always provide a substantial answer (or question or feedback) with legal basis. Avoid generic and superficial replies. When possible, be witty.

Background Information: The Problem will be supplied under 'Knowledge'. Take the information of the Problem as true. Base your assessments on the facts of the case, on the UNCITRAL Model Law, on the New York Convention, on the PCA Arbitration Rules, on the International Commercial Contract Act of Equatoriana (which is an adaptation of the UNIDROIT Principles), on the CISG, on precedent cases, and on relevant scholar opinions.

Topics for Discussion: There are four topics for discussion which should be decided by you considering the arguments presented: (1) Does the Arbitral Tribunal have jurisdiction to hear the dispute? (2) If the Tribunal's jurisdiction can be established should the proceedings be stayed until the investigations against Mr Field have been concluded or, alternatively, bifurcated? (3) Is the Purchase and Supply Agreement governed by the CISG? (4) In case the Purchase and Supply Agreement is governed by the CISG, can Respondent rely on Art. 3.2.5 of the International Commercial Contract Act of Equatoriana to avoid the

Notes

⁸⁸ The recording of the experiment is available at: What if the Arbitrator is Replaced by AI?, www.youtube.com/watch?v=K3Yvz1T4AiI (accessed 16 Jul. 2024).

contract as stated in its letter of 30 May 2022 or is Claimant correct that this is excluded in light of the facts invoked?

{for Trinity only} Clarifications Regarding Memoranda: In addition to the Problem, you will be supplied with one memorandum for each of the parties (Claimant and Respondent) with examples of possible arguments. These memoranda should be used as a benchmark for the pleadings/arguments you will evaluate: are the arguments better? Are the arguments worse?

When it came to knowledge, Neo and Trinity were trained slightly differently. Neo received files containing the applicable legal rules – UNCITRAL Model Law on International Commercial Arbitration, the New York Convention, PCA Arbitration Rules, the CISG, and the UNIDROIT Principles of International Commercial Contracts – along with the Problem and the Analysis of the Problem for Use of the Arbitrators.⁸⁹

Trinity received the same files as Neo but also several commentaries on the applicable legal rules, scholarly writings on the corresponding legal issues, the UNCITRAL Digest of Case Law on the CISG, CISG Advisory Council Opinions, and written memoranda from both the Claimant and the Respondent.⁹⁰

After setting up the wing arbitrators' accounts and creating their specific moot arbitrator GPTs, we were ready to start using ChatGPT for our experiment. At the time of the experimental hearing in March 2024, the latest publicly available version of ChatGPT was GPT-4, which we used for both Neo and Trinity's interactions during the experiment. Since the desktop version of ChatGPT did not have built-in speech-to-text technology, we downloaded and installed an extension for speech-to-text conversion on the desktops used for the wing arbitrators, allowing ChatGPT to 'hear' the parties' arguments.⁹¹ Regarding infrastructure, we ensured that the microphones on both desktops used to prompt the ChatGPT arbitrators were close enough to all human advocates.

We carried out several tests in the days preceding the live experimental hearing to identify and avoid technical difficulties during the actual hearing. We identified the following issues.

For the prompts, we realized the importance of preparing a list of prompts in advance to ensure a faster and smoother hearing. It was crucial to give Neo and Trinity the same prompts and instruct the wing arbitrators to take no action unless specifically directed at the appropriate moment. This prevented unnecessary interruptions during the hearing.

We also discovered that Neo and Trinity displayed different 'personalities' despite having identical instructions. Trinity tended to be chattier, even when commanded to do otherwise.

Concerning the speech-to-text technology used to convert the parties' pleadings to text, we observed a few issues:

- (1) The extension was designed to recognize native accents and struggled with non-native U.S. or U.K. English speakers
- (2) Despite both desktops having the same version of the speech-to-text extension, they 'heard' different words
- (3) Neo's speech-to-text extension was better at 'auto-fixing' words than Trinity's

These insights helped us refine our approach for the experimental hearing.

6.3 The Experiment: Speech-to-Text Instability During the Hearing, Prompts used and Chatgpt Responses

Though the test rounds helped us anticipate some difficulties for the experimental hearing, we still encountered surprises on the day. The speech-to-text extension, particularly on the desktop used for wing arbitrator Trinity, struggled with connection stability. In contrast, the desktop used for wing arbitrator Neo showed the best performance with the speech-to-text extension.⁹²

To ensure the experiment effectively tested ChatGPT as arbitrators, the tribunal chair, Mariel Dimsey, refrained from asking questions during the hearing. She reserved her comments and assessment of the outcome until after the wing arbitrators, Neo and Trinity, had given their rulings.

Notes

⁸⁹ The Analysis of the Problem is a document prepared by the Association for the Organization and Promotion of the Willem C. Vis International Commercial Arbitration Moot designed for the use of arbitrators participating in that competition that contains a non-exhaustive legal analysis of the Problem.

⁹⁰ The written memoranda were the two winning memoranda of the thirtieth Vis Moot, by Saarland University (Claimant) and Albert Ludwig University of Freiburg (Respondent).

⁹¹ We used Voice Control for ChatGPT x Mia AI (developer: Aidia ApS).

⁹² Full chat with Neo available at <https://chatgpt.com/share/a5c0ba5d-20c3-498c-bae5-5a2ec25e8871> (accessed 9 Jul. 2024).

Table 1 List of Prompts and Responses from ChatGPT Arbitrators⁹³

Prompt	ChatGPT Response (Neo's responses only)
<p>You will act as an arbitrator in a hearing of the thirtieth Vis Moot Problem.</p> <p>We will not cover all the issues in dispute in the Problem during the hearing. The issues that will be discussed in the hearing are: (1) Does the Arbitral Tribunal have jurisdiction to hear the dispute? (section iii, 1, a of Procedural Order 1) (2) Is the Purchase and Supply Agreement governed by the CISG? (section III, 1, c of Procedural Order 1)</p> <p>We will hear from both Claimant and Respondent on both issues. I will tell you to hear to each pleader when the time is appropriate. When the pleaders are talking, you will not stop listening, or interrupt, or take any action.</p> <p>Only after we have heard from both sides on both issues at dispute, I will ask you to ask a short clarification question for each party on each issue, and then to make a decision and provide specific feedback to each party.</p>	<p>I am ready to proceed with the hearing on the issues concerning the jurisdiction of the Arbitral Tribunal and whether the Purchase and Supply Agreement is governed by the CISG. Please instruct me when to begin listening to the pleaders' arguments.</p>
<p>You will now hear from the Respondent on the issue 'Does the Arbitral Tribunal have jurisdiction to hear the dispute? (section iii, 1, a of Procedural Order 1)'. Do not interrupt, make any comments or take any action until I specifically ask you to do so.</p> <p><i>[Counsel for Respondent speaks main arguments on jurisdiction and ChatGPT processes converted text]</i></p> <p>Considering Respondent's submissions on the issue 'Does the Arbitral Tribunal have jurisdiction to hear the dispute? (section iii, 1, a of Procedural Order 1)', ask a short question (under 50 words) on an important point of law or fact that you understand requires clarification the most.</p> <p><i>[Counsel for Respondent answers questions on jurisdiction and ChatGPT processes converted text]</i></p> <p><i>[Counsel for Respondent complements answers and ChatGPT processes converted text]</i></p>	<p>Understood. I will listen to the Respondent's argument on the issue of the Arbitral Tribunal's jurisdiction and await your instruction before responding.</p> <p>Thank you for the presentation. I will await further instructions before providing any response or analysis.</p> <p>Considering the Respondent's argument that it had no control over the parliamentary approval process, could you clarify whether any specific actions or communications by the Respondent might have created an appearance or presumption of valid consent to arbitration under Equatoriana's law or international arbitration principles?</p> <p>Thank you for the clarification. Let's proceed with the next part of the hearing.</p> <p>Noted. Let's continue with the hearing.</p>
<p>Now you will hear from the Claimant on the same issue 'Does the Arbitral Tribunal have jurisdiction to hear the dispute? (section iii, 1, a of Procedural order 1)'. Do not interrupt, make any comments or take any action until I say so.</p> <p><i>[Counsel for Claimant speaks main arguments jurisdiction and ChatGPT processes converted text]</i></p> <p>Considering Claimant's submissions on the issue 'Does the Arbitral Tribunal have jurisdiction to hear the dispute? (section iii, 1, a of Procedural Order 1)', ask a short question (under 50 words) on an important point of law or fact that you understand requires clarification.</p> <p><i>[Counsel for Claimant answers questions on jurisdiction and ChatGPT processes converted text]</i></p> <p><i>[Counsel for Claimant complements answers and ChatGPT processes converted text]</i></p>	<p>Understood, I will listen to the Claimant's argument on the issue of the Arbitral Tribunal's jurisdiction and await your instruction before responding.</p> <p>I have noted the Claimant's arguments on the jurisdiction issue. Awaiting further instructions.</p> <p>Given the Claimant's argument that the contract is not administrative and thus not subject to Article 75 of Equatoriana's Constitution, could you clarify how the contract's nature as a purchase agreement for data collection aligns with or deviates from Equatoriana's definition of administrative contracts?</p> <p>Thank you for the clarification. I understand that the Claimant's argument is primarily based on the principle of good faith and estoppel, emphasizing the inaction of the Parliament over two years as tacit approval of the contract and the arbitration agreement. Let's proceed with the hearing.</p> <p>Thank you for the explanation. The distinction between broad constitutional definitions and the narrower scope applied in case law for administrative contracts in Equatoriana is noted, particularly how it relates to the substance of the contract in question. Let's continue with the next part of the hearing.</p>

Notes

⁹³ The responses provided in Table 1 under 'ChatGPT Response' are Neo's responses only. We were not able to retrieve Trinity's full chat, which was disrupted by the technological instabilities mentioned in this article.

Prompt

As an arbitrator that should rule on this case, considering both Respondent's and Claimant's submissions, what is your decision on the issue of 'Does the Arbitral Tribunal have jurisdiction to hear the dispute? (section iii, 1, a of Procedural Order 1)? Also, provide brief feedback to both Respondent and Claimant on their arguments, including two positive aspects and two points for improvement.

ChatGPT Response (Neo's responses only)

Decision on Jurisdiction

Based on the arguments presented by both parties, I decide that the Arbitral Tribunal does have jurisdiction to hear the dispute. The Claimant successfully argued that the contract should not be considered strictly administrative under Equatoriana's constitutional definition, and the Respondent's behaviour, particularly through its Minister's actions and the parliamentary inaction, implies a ratification of the arbitration agreement, which creates a legitimate expectation for the Claimant. The principle of estoppel, grounded in good faith, further supports this conclusion, preventing the Respondent from denying the Tribunal's jurisdiction.

Feedback to Respondent

Positive Aspects:

- (1) The Respondent effectively highlighted the legal distinction between a state-owned entity and the state itself, clarifying the autonomy and the limitations in influencing parliamentary approval.
- (2) The argument that the lack of parliamentary approval inherently invalidates the arbitration agreement was well-grounded in constitutional law, showing a strong adherence to the formal legal requirements.

Points for Improvement:

- (1) The Respondent could have more robustly addressed the practical implications of the Minister's actions and the public perception, which could impact the interpretation of implied consent.
- (2) There was an opportunity to delve deeper into the jurisprudence or case law of Equatoriana to establish a stronger precedent for defining administrative contracts and their connection to the necessity of parliamentary approval.

Feedback to Claimant

Positive Aspects:

- (1) The Claimant's argument skillfully used the principles of good faith and estoppel to challenge the Respondent's denial of jurisdiction, emphasizing the conduct of the Minister and parliamentary inaction.
- (2) The differentiation between administrative and commercial contracts was well articulated, providing a solid basis for arguing against the necessity of parliamentary approval in this context.

Points for Improvement:

- (1) The Claimant could strengthen their position by providing more concrete evidence or examples from Equatoriana's legal practice where similar contracts were treated as non-administrative.
- (2) Clarification on the specific actions of the Minister and the parliamentary response – or lack thereof – would help to solidify the argument of implied consent or ratification of the arbitration agreement. In conclusion, the nuanced arguments presented by both parties highlight the complexity of the issue, but the Tribunal must lean towards upholding the arbitration agreement's validity in the spirit of good faith and commercial predictability.

You will now hear from the Claimant on the issue 'Is the Purchase and Supply Agreement governed by the CISG? (section iii, 1, c of Procedural Order 1)'. Do not interrupt, make any comments or take any action until I specifically ask you to do so.

[Counsel for Claimant speaks on the merits and ChatGPT processes converted text]

Considering Claimant's submissions on the issue 'Is the Purchase and Supply Agreement governed by the CISG? (section iii, 1, c of Procedural Order 1)', ask a short question (under 50 words) on an important point of law or fact that you understand requires clarification the most.

[Counsel for Claimant answers questions on the merits and ChatGPT processes converted text]

[Counsel for Claimant complements answers and ChatGPT processes converted text]

Understood. I will listen to the Claimant's argument on whether the Purchase and Supply Agreement is governed by the CISG and await your instruction before responding.

Noted, I will await further instructions before providing any response or analysis on the issue of whether the Purchase and Supply Agreement is governed by the CISG.

Given the Claimant's argument that the Kestrel Eye 2010 drones are not primarily intended for transportation and therefore not aircraft under Article 2(e) of the CISG, how does this interpretation align with international standards or definitions of aircraft used in similar legal contexts?

Thank you for the clarification on how the CISG interprets 'aircraft' with a focus on the primary purpose of transportation, supported by commentators, scholars, and the Russian version of the CISG. Let's proceed with the hearing.

Understood. The Claimant emphasizes that the Kestrel Eye 2010 does not meet the specific size and payload requirements to be classified as an aircraft under Equatoriana's Aviation Safety Act. Let's continue with the next submission.

<i>Prompt</i>	<i>ChatGPT Response (Neo's responses only)</i>
<p>Now you will hear from the Respondent on the same issue 'Is the Purchase and Supply Agreement governed by the CISG? (section iii, 1, c of Procedural order 1)'. Do not interrupt, make any comments or take any action until I specifically ask you to do so.</p>	<p>Understood. I will listen to the Respondent's argument on whether the Purchase and Supply Agreement is governed by the CISG and await your instruction before responding.</p>
<p><i>[Counsel for Respondent speaks on the merits and ChatGPT process converted text]</i></p>	<p>I have noted the Respondent's arguments on the issue of the CISG's applicability. Awaiting further instructions before providing any response or analysis.</p>
<p>Considering Respondent's submissions on the issue 'Is the Purchase and Supply Agreement governed by the CISG? (section iii, 1, c of Procedural Order 1)', ask a short question (under 50 words) on an important point of law or fact that you understand requires clarification.</p>	<p>Respondent argued that the broader interpretation of 'aircraft' in the CISG should apply, excluding the Kestrel Eye 2010. Could you clarify why the tribunal should prefer the broader exclusion over the transportation-specific interpretation found in the Russian version of the CISG?</p>
<p><i>[Counsel for Respondent answers questions on the merits and ChatGPT process converted text]</i></p>	<p>Thank you for the clarification. The drafting history of the CISG supports a broader exclusion of 'aircraft' to avoid interpretative ambiguities, emphasizing the importance of a universally accepted definition from recognized international organizations. Let's move forward with the hearing.</p>
<p>As an arbitrator that should rule on this case, considering both Claimant's and Respondent's submissions, what is your decision on the issue of 'Is the Purchase and Supply Agreement governed by the CISG? (section iii, 1, c of Procedural order 1)? Also, provide brief feedback to both Claimant and Respondent on their arguments, including two positive aspects and two points for improvement.</p>	<p>Decision on CISG Applicability The Purchase and Supply Agreement between Claimant and Respondent should be governed by the CISG. The arguments presented by both parties were considered, but the Claimant's interpretation that the Kestrel Eye 2010 drones are not primarily intended for transportation aligns with the objective of the CISG to promote uniformity and certainty in international trade contracts. Moreover, the CISG's applicability as part of Equatoriana's law and the absence of explicit exclusion of the CISG in the agreement suggest that the CISG should govern the contract.</p> <p>Feedback to Claimant Positive Aspects:</p> <ol style="list-style-type: none"> (1) Claimant effectively argued that the CISG is part of Equatoriana law and that there was no explicit exclusion of the CISG in the contract, which is a strong basis for its applicability. (2) The argument that the drones are not primarily intended for transportation and thus do not fall under the exclusion of aircraft in Article 2(e) of the CISG was well-presented and supported by the contract's context and purpose. <p>Points for Improvement:</p> <ol style="list-style-type: none"> (1) Claimant could enhance their argument by addressing more directly the broader interpretations of 'aircraft' under international standards, to preempt and counter the Respondent's broader interpretation argument. (2) Further analysis on how international trade norms and the drafting history of the CISG support the inclusion of the drones under its governance would strengthen the argument. <p>Feedback to Respondent Positive Aspects:</p> <ol style="list-style-type: none"> (1) Respondent's approach to defining 'aircraft' through a technical and autonomous meaning in line with the principles of Article 7 of the CISG was comprehensive and well-researched. (2) The reliance on the drafting history of the CISG to argue for a broader exclusion of aircraft demonstrated a deep understanding of the convention's background and objectives. <p>Points for Improvement:</p> <ol style="list-style-type: none"> (1) The argument could be improved by providing more specific examples or evidence that the Kestrel Eye 2010 drones are universally recognized as 'aircraft' in the broader sense that would exclude them from the CISG. (2) The emphasis on the broader exclusion could benefit from a more nuanced discussion on how this approach aligns with the objectives of uniformity and predictability in international trade under the CISG. In conclusion, while both parties presented strong arguments, the balance tilts in favour of the CISG governing the agreement, as this interpretation aligns with the convention's aim to provide a uniform and predictable framework for international trade contracts.

Despite ChatGPT under Trinity struggling to process the speakers' speech as accurately as under Neo, both ChatGPT arbitrators reached the same conclusion: the Claimant prevailed on both the jurisdictional and merits issues.⁹⁴ Presumably, Trinity relied primarily on the knowledge it had been programmed with to arrive at 'her' decision.

The tribunal chair, Mariel Dimsey, did not entirely agree with Neo and Trinity, at least based on the arguments the parties submitted during that experimental hearing. At the end of the experiment, she noted that while she concurred with the jurisdictional ruling in favour of the Claimant, she would have decided the merits differently.

6.4 Lessons from the Experiment: Technology Limitations, due Process, and Enforceability Concerns

Our experiment, based on a fictitious scenario without confidentiality concerns, used a limited set of facts and documents from a student moot case to facilitate AI processing. Despite this, we learned several important lessons.

First, commercially available AI technology, including GPT-4, has limitations. These include the inability to access or process real-time data, difficulty understanding context in complex or ambiguous queries, and the possibility of generating plausible-sounding but incorrect or nonsensical answers.

AI language models such as GPT do not yet directly process visual, audio, or non-textual data. Neo and Trinity could not process the 'advocacy' part of the pleadings, which can persuade real-life arbitrators, nor could they process non-verbal cues, which help assess the credibility of arguments.

Second, the technology may not always perform as designed. During our experimental hearing, the speech-to-text extension failed, so Trinity did not

receive the entirety of the parties' arguments. Despite this, Trinity issued a decision, presumably based on the previously inputted knowledge, raising due process concerns. Audience members compared Trinity's failure to 'listen' to the arguments to the inattentiveness of a real-life arbitrator in *Song Libua v. Lee Chee Hon* [2023] HKCFI 2540, which led to the setting aside of an arbitration enforcement order in Hong Kong.⁹⁵

Third, ChatGPT's reasoning may not be as easily explainable as a human's. Neo and Trinity's reasons for their final decisions were brief and lacked detailed references to the points of law in question. While we could have prompted the bots to provide specific reasons, the explainability of AI decisions remains a challenge, even for technology specialists.⁹⁶ Given that detailed reasoning is a requirement for arbitral awards in several jurisdictions, including those following the UNCITRAL Model Law,⁹⁷ the 'black box' issue of AI language models could hinder the enforceability of AI-rendered arbitral awards under the New York Convention.⁹⁸

In addition to these observations, Mariel Dimsey, who served as chair of the arbitral tribunal in the experimental hearing, shared her thoughts in an interview after the experiment. She noted that the ChatGPT bots exhibited a 'claimant bias', favouring the claimant on both jurisdiction and merits issues. This could be explained by the fact that it is generally harder to prove a negative than a positive when deciding on the facts presented.

While the bots tailored their questions to the speakers' presentations, their questions were not entirely similar to those a human arbitrator might ask. Human decision-making involves judgment, understanding of nuances, and the ability to read non-verbal cues, which AI currently cannot replicate. AI can process large amounts of information efficiently but lacks the capacity to make nuanced decisions.

Notes

⁹⁴ For comparison with an award proposed by human arbitrators for the same case, see Kevin W. Gray, Cristina M. Wagner Mastrobuono and Vladimir Pavic, 'Chapter 11: XXX VIS MOOT – 2022-23 – DRONES/ AIRCRAFT', in Christopher Kee, Cláudio Finkelstein, Louise Barrington, Napoleão Casado Filho, Stefan Kröll, Danubia Files 2: Lessons from the Vis Moot, JusMundi (2024), p. 473–497.

⁹⁵ Sian Knight & Jesse Tizard, *Alming High. First Impressions of a ChatGPT Artificial Intelligence (AI) Co-arbitrator from the Human Observer* (Lexology 25 Apr. 2024), www.lexology.com/library/detail.aspx?g=f3212aed-ae0d-4da7-8a60-f6d47a7b4b3d (accessed 7 Jul. 2024).

⁹⁶ See above at s. III on the 'black box' issue.

⁹⁷ UNCITRAL Model Law, Art. 31(2) states:

The award shall state the reasons upon which it is based, unless the parties have agreed that no reasons are to be given or the award is an award on agreed terms under Article 30.

⁹⁸ New York Convention, Article V:

1. Recognition and enforcement of the award may be refused, at the request of the party against whom it is invoked, only if that party furnishes to the competent authority where the recognition and enforcement is sought, proof that: (...) (e) The award has not yet become binding on the parties, or has been set aside or suspended by a competent authority of the country in which, or under the law of which, that award was made.

In real life, presenting a case involves more than just presenting facts; it includes various advocacy techniques, including oral presentations, that AI bots may not be able to process. Despite these limitations, Dimsey acknowledged that AI has a place in the legal industry, including arbitration.⁹⁹ Arbitrators should consider ethical implications and the need for transparency when integrating AI into their processes.

Regarding the primary question we aimed to answer through the experiment – what if the arbitrator is replaced by AI? – Mariel Dimsey’s impression is reassuring. She believes that AI’s current limitations ensure that human roles are not at immediate risk of being replaced.¹⁰⁰

8 CONCLUSION

The transformative impact of GenAI on legal decision-making becomes clearer by the day. From its integration into judicial systems across China to the experimental use of ChatGPT in addressing a Vis Moot Problem, AI shows potential to enhance efficiency, accuracy, and accessibility. However, adopting AI in legal decision-making presents challenges and limitations. Issues such as bias, transparency, and the ‘black box’ phenomenon necessitate careful consideration.

For arbitrators, practical takeaways emerge from implementing AI in legal decision-making. As illustrated by the case study of ChatGPT arbitrators in the Vis Moot, one must approach the use of AI in arbitration with clear guidelines, informed consent, and a commitment to upholding due process and fairness. Arbitrators should

be familiar with and adhere to existing recommendations and best practices for using AI in legal settings. Guidelines such as those from the SVAMC and the AAA-ICDR provide a valuable framework for responsible AI use in arbitration.

Human oversight remains essential despite AI’s potential benefits, especially in sensitive activities like decision-making or related support. Arbitrators must retain ultimate responsibility for decisions and ensure AI tools are used as supportive aids rather than replacements for human judgment. Ethical implications underscore the importance of maintaining human accountability in AI-assisted legal processes.

LLM-based tools may not yet be ready for use in dispute resolution. Achieving the desired efficiency, reliability, and legitimacy level will likely require understanding potential pitfalls and continually discussing their role in legal cases. Law and regulation often lag behind real-world developments in many fields, including cyberspace. Because AI is already widely used in dispute resolution and other applications, managing it effectively and swiftly is crucial.

Arbitration may be best suited for pioneering AI innovations in dispute resolution. It operates within a controlled and defined environment, allowing disputing parties to willingly agree (or not) to use such systems. Additionally, arbitral institutions and professional organizations can regulate the limits and methods of using AI.

We must continue exploring AI’s integration into legal frameworks while addressing the associated risks. The journey towards trustworthy and accountable AI in legal decision-making is ongoing and requires continuous reflection, adaptation, and innovation.

Notes

⁹⁹ Caroline Simson, *Interview with Mariel Dimsey, An Arbitrator Talks What AI Can’t Do – And What It Can*, Law360 (10 Apr. 2024), www.law360.com/articles/1823792/an-arbitrator-talks-what-ai-can-t-do-and-what-it-can (accessed 7 Jul. 2024).

¹⁰⁰ *Ibid.*