

**** Disclaimer: these Guidelines have been made publicly available in draft form for the purposes of receiving feedback and comments, and should not be used, adapted or relied on before the final version has been published by SVAMC.*

SILICON VALLEY ARBITRATION AND MEDIATION CENTER

**GUIDELINES ON THE USE OF ARTIFICIAL INTELLIGENCE
IN ARBITRATION**

DRAFT of 31 August 2023

INTRODUCTION

These Guidelines on the Use of Artificial Intelligence in Arbitration (the **Guidelines**) introduce a principle-based framework for the use of artificial intelligence (**AI**) tools in arbitration at a time when such technologies are becoming increasingly powerful and popular. They are intended to assist participants in arbitrations with navigating the potential applications of AI.

These Guidelines can be used in domestic or international arbitrations and are meant to serve as a point of reference for arbitral institutions, arbitrators, parties and their representatives (including counsel), experts, and, where relevant, other participants in the arbitral process. To that end, the Guidelines provide a Model Clause that can be incorporated into procedural orders to make the Guidelines applicable to all participants involved in a particular arbitration proceeding.

The Guidelines are prefaced by preliminary provisions which clarify the scope and application of the principles contained herein. The body of the Guidelines is organised into three chapters: one chapter containing Guidelines that generally apply to all participants in the arbitration process, regardless of their role; a second chapter containing Guidelines that address specific uses of AI by parties and party representatives (including counsel); and a third chapter with Guidelines addressing particular considerations that may arise when arbitrators use AI.

In a separate section, the Guidelines offer examples of both compliant and non-compliant uses of AI in arbitrations. These examples are illustrative only to clarify the practical implications of the Guidelines and provide a yardstick to measure conformity in real-world scenarios.

Members of the SVAMC AI Task Force Guidelines Drafting Subcommittee

Orlando Federico Cabrera Colorado, Elizabeth Chan, Dmitri Evseev, Marta García Bel, Sofia Klot, Benjamin I. Malek, Soham Panchamiya and Duncan Pickard.

**** Disclaimer: these Guidelines have been made publicly available in draft form for the purposes of receiving feedback and comments, and should not be used, adapted or relied on before the final version has been published by SVAMC.*

SUMMARY OF GUIDELINES

PRELIMINARY PROVISIONS

Application of the Guidelines

Definition of AI

Non-derogation of any mandatory rules

CHAPTER 1: Guidelines applicable to all participants in international arbitrations

1. Understanding the uses, limitations, and risks of AI applications
2. Safeguarding confidentiality
3. Disclosure and protection of records

CHAPTER 2: Guidelines for parties and party representatives

4. Duty of competence or diligence in the use of AI
5. Respect for the integrity of the proceedings and evidence

CHAPTER 3: Guidelines for arbitrators

6. Non-delegation of decision-making responsibilities
7. Respect for due process

EXAMPLES OF COMPLIANT AND NON-COMPLIANT USES OF AI IN ARBITRATIONS

MODEL CLAUSE FOR INCLUSION IN PROCEDURAL ORDERS

*** Disclaimer: these Guidelines have been made publicly available in draft form for the purposes of receiving feedback and comments, and should not be used, adapted or relied on before the final version has been published by SVAMC.

PRELIMINARY PROVISIONS

Application of the Guidelines

These Guidelines shall apply when and to the extent that the parties have so agreed and/or following a decision by an arbitral tribunal or an arbitral institution to adopt these Guidelines.

Commentary

The Guidelines seek to establish a set of general principles for the use of AI in arbitration. Intended to guide rather than dictate, they are meant to accommodate case-specific circumstances and technological developments, promoting fairness, efficiency, and transparency in arbitral proceedings.

These Guidelines may be adopted, in whole or in part, in the arbitration agreement or by the parties and/or the tribunal at any other time subsequently, including during the course of arbitral proceedings (*see* Model Clause for inclusion in Procedural Orders).

As applied to international arbitrations, the Guidelines acknowledge the multi-faceted and multi-jurisdictional nature of such proceedings. Given the potential for various national laws to apply –for instance, an arbitration seated in Paris, governed by Mexican law, with hearings in Hong Kong– it becomes necessary to harmonise the potentially disparate local and international standards relating to the use of AI.

Accordingly, these Guidelines do not intend to replace or override local AI laws or regulations (*see* non-derogation of any mandatory rules). Instead, they serve as a supplementary international standard that provides a common denominator for AI’s ethical and effective use in international arbitrations.

Development of best practices around the use of AI in international arbitration is only beginning, and these Guidelines aim to contribute to that effort. As such, they are a tool that assists parties, arbitral tribunals, institutions and others in navigating the application of AI, with an understanding that technologies, local laws and international standards will continue to evolve.

Definition of AI

As used in these Guidelines, the term “AI” refers to computer systems that perform tasks commonly associated with human cognition, such as understanding natural language, recognising complex semantic patterns, and generating human-like outputs.

**** Disclaimer: these Guidelines have been made publicly available in draft form for the purposes of receiving feedback and comments, and should not be used, adapted or relied on before the final version has been published by SVAMC.*

Commentary

There is no single definition of AI, and even existing definitions may evolve over time. For this reason, it is essential to clarify how the term should be understood in the Guidelines.

The definition adopted is meant to be broad enough to encompass both existing and future foreseeable types of AI but not so broad as to encompass every type of computer-assisted automation tool. Rather, the definition focuses on modern technologies that tend to be more autonomous, complex, multifunctional and probabilistic than traditional automation tools based on rule-based deterministic logic.

Modern AI systems are usually based on machine learning, a set of computer science techniques that allow machines to learn patterns and make intelligent predictions based on the data on which they have been trained. Machine learning algorithms have existed for decades and are employed behind the scenes in various technology products used by dispute resolution professionals, such as spelling and grammar checkers, email spam filters, search engines, optical character recognition (also known as “OCR”), and machine translation.

With the advent of technological advances such as deep neural networks, large language models and generative AI, however, it has become possible for the general public to interact with multi-purpose AI systems directly. The potential uses for AI in the field of dispute resolution has exploded, even as the risks and limitations of these tools have become more difficult to comprehend. For example, deep neural networks can learn highly complex patterns and abstractions. Still, these patterns are recorded in a largely indecipherable form even to the computer scientists who created the models. Moreover, such models generate outputs based on statistical probabilities rather than a defined set of rules.

Large language models are a type of deep neural network trained on vast amounts of textual data and capable of generating natural-sounding and plausible (but not necessarily accurate) responses to a given prompt. AI that can generate meaningful text, images or other types of output that appears creative and extrapolates well beyond the data the model was trained on is often referred to as Generative AI. Generative AI is used in tasks such as question-answering, summarising text and producing drafts based on a given input or instruction.

It is important to note that, while Generative AI systems tend to receive the most publicity and are the most accessible to the general public, there are other equally complex types of AI, such as those powering recommendation or classification tools, sometimes known as evaluative or discriminative AI. The focus of these Guidelines is not solely on Generative AI but rather on all modern types of AI tools, whether intended to perform a specific evaluation or to generate outputs that resemble human-created content (including text, sound or visual images).

**** Disclaimer: these Guidelines have been made publicly available in draft form for the purposes of receiving feedback and comments, and should not be used, adapted or relied on before the final version has been published by SVAMC.*

Non-derogation of any mandatory rules

These Guidelines shall not derogate from any legal obligations, ethical duties, or rules of professional conduct, or any other binding rules applicable to the arbitration proceedings or persons participating in them.

Commentary

This provision recognises that the use of AI tools and AI applications in arbitrations may be subject to a range of rules and regulations, whether at the domestic or international level. These include, but are not limited to, laws, domestic statutes or international treaties on the use and development of AI, domestic rules of professional conduct, ethical and professional standards, and applicable arbitration rules, all of which can indirectly impact how certain professionals can use AI tools in an arbitration setting.

These Guidelines should not be construed as detracting or derogating from any of the above-mentioned rules and regulations. To the extent that these Guidelines are incompatible with any applicable mandatory rules and regulations, the latter should prevail.

*** Disclaimer: these Guidelines have been made publicly available in draft form for the purposes of receiving feedback and comments, and should not be used, adapted or relied on before the final version has been published by SVAMC.

CHAPTER 1: GUIDELINES FOR ALL PARTICIPANTS IN INTERNATIONAL ARBITRATIONS

GUIDELINE 1

Understanding the uses, limitations, and risks of AI applications

All participants involved in arbitration proceedings who use AI tools in preparation for or during an arbitration are responsible for familiarising themselves with the AI tool’s intended uses and should adapt their use accordingly.

All participants using AI tools in connection with an arbitration should make reasonable efforts to understand each AI tool’s relevant limitations, biases, and risks and, to the extent possible, mitigate them.

Commentary

Participants should make reasonable efforts to understand, at least in general terms, the functionality, limitations and risks of the AI tools they use in preparation for or during the course of an arbitration proceeding. For example, for tools that use Generative AI, participants should recognise the known limitations of such tools, such as their tendency to perpetuate biases contained in the training data, their propensity to mix up or invent information to fill gaps in knowledge, and their inability to identify the true logic or sources of information used to produce a given output, as further described below.

Participants should also review the AI tool’s terms of use and data handling policies to understand if the tool’s data treatment is consistent with any applicable confidentiality, privacy, or data security obligations (*see* Guideline 2 - Safeguarding confidentiality).

Notably, participants should be aware of the following limitations, biases, and risks that (at present) are inherent in the use of certain AI tools:

“Black-box” problem

Generative AI tools produce natural-sounding and contextually relevant text based on speech patterns and semantic abstractions learned during their training. However, these outputs are a product of infinitely complex probabilistic calculations rather than intelligible “reasoning” (the so-called “black box” problem). Despite any appearance otherwise, AI tools lack self-awareness or the ability to explain their own algorithms.

**** Disclaimer: these Guidelines have been made publicly available in draft form for the purposes of receiving feedback and comments, and should not be used, adapted or relied on before the final version has been published by SVAMC.*

In response to this problem, participants may, as far as practical, use AI tools and applications that incorporate explainable AI features or otherwise allow them to understand how a particular output was generated based on specific inputs. “Explainable AI” is a set of processes and methods that allows human users to comprehend how an AI system arrives at a certain output based on specific inputs. Explainable AI can help promote transparency, increase trust in the AI tool’s accuracy and help ensure fairness when applied in an arbitration context, especially when the output of an AI tool significantly influences the proceedings. However, a complete understanding of complex AI systems may be beyond the reach of non-technical individuals, and this Guideline does not impose an expectation of thorough understanding. There are also technical and cost-related limitations to explain how AI systems work fully, especially those systems employing complex algorithms and machine learning techniques.

Quality and representativeness of the training data

Large language models and other AI tools are trained using specific datasets and parameters, and their capabilities are a function of that particular training. Even the most advanced AI tools will exhibit biases and blind spots resulting from limitations in underlying datasets and training protocols. Moreover, general-purpose AI tools may not be well-suited for tasks requiring specialised knowledge or case-specific information, unless they are fine-tuned or provided with more relevant data.

Errors or “hallucinations”

Large language models have a tendency to “hallucinate” or offer incorrect but plausible-sounding responses when they lack information to provide an accurate response to a particular query. Hallucinations occur because these models use mathematical probabilities (derived from linguistic and semantic patterns in their training data) to generate a fluent and coherent response to any question. However, they typically cannot assess the accuracy of the resulting output.

Hallucinations can be reduced through various techniques such as “prompt engineering” (*i.e.* crafting the query in a manner that is more likely to generate a better response) and “retrieval-augmented generation” (*i.e.* providing the model with relevant source material together with the query), but they are difficult to eliminate completely.

Augmentation of biases

An AI tool’s training data may reflect biases that can be perpetuated through the use of the tool. Participants in arbitrations should minimise the risks associated with flawed or biased predictions by exercising their own independent judgement.

This is especially important when existing biases in the data may create, exacerbate or perpetuate any form of discrimination, racial, gender or other profiling in the search and appointment of individuals as arbitrators, experts, counsel, or any other roles in connection with arbitrations. Biases may occur when the underrepresentation of certain groups of individuals is carried over to the training data used by the AI tool to make selections or assessments. Participants should exercise extreme caution in using any AI tool for such purposes, especially if they are unaware of how the selection or assessment algorithm works.

**** Disclaimer: these Guidelines have been made publicly available in draft form for the purposes of receiving feedback and comments, and should not be used, adapted or relied on before the final version has been published by SVAMC.*

Using AI tools to help identify a suitable candidate for a specific role in connection with an arbitration is a particularly sensitive matter, and participants should be mindful of the impact such use may have on diversity and the fair representation of diverse individuals.¹ In summary, participants are urged to: (i) use their personal judgement to evaluate the output of these AI tools from a diversity standpoint; (ii) to the best of their ability, become aware of the potential biases that may underlie the AI tool's output and, to the extent possible, mitigate them; (iii) use AI tools that control for biases.

¹ The term “diversity”, as used in this Commentary, refers to race, ethnicity, national origin, gender, sexual orientation, gender identity and ability.

**** Disclaimer: these Guidelines have been made publicly available in draft form for the purposes of receiving feedback and comments, and should not be used, adapted or relied on before the final version has been published by SVAMC.*

GUIDELINE 2

Safeguarding confidentiality

All participants in international arbitration are responsible for ensuring their use of AI tools is consistent with their obligations to safeguard confidential information (including privileged, private, secret or otherwise protected data). They should not submit confidential information to any AI tool without appropriate vetting and authorisation. Where a third-party (rather than an in-house) AI tool is considered for use in arbitration, special attention should be paid to the third party's policies on recording, storage and use of prompt or output histories and of any other confidential data sources provided to the AI tool.

Only AI tools that adequately safeguard confidentiality should be approved for uses that involve sharing confidential or legally privileged information with third parties. For this purpose, participants should review the data use and retention policies offered by the relevant AI tools and opt for more secure solutions.

Parties and their representatives should be aware of the data and confidentiality risks associated with using particular AI tools available to the general public in connection with an arbitration.

Commentary

Different jurisdictions have their own rules on confidentiality, privilege and secrecy of information.

Professionals bound by these duties should limit themselves to using AI tools that adequately safeguard the confidentiality of client or other protected data, or otherwise refrain from inputting any such data into AI tools that do not guarantee confidentiality.

Some AI tools available to the general public may retain information provided to them for a variety of purposes or even state that the service providers have rights to all the information that users enter. The use of these publicly available AI tools in the context of an arbitration could pose a risk of disclosing confidential information. By contrast, business-oriented or privacy-oriented AI tools and vendors may offer similar functionality but with additional safeguards for confidentiality.

GUIDELINE 3

Disclosure and protection of records

[Note for reviewers: The Drafting Subcommittee has produced two alternative drafts of most provisions of this Guideline for consideration. The key substantive differences between Options A and B are highlighted for convenience in the PDF. In your comments, we kindly ask that you indicate a preference for Option A or Option B along with any other comments or suggestions you wish to make.]

[OPTION A]

1-1. Disclosure concerning the use of AI tools in connection with an arbitration may be appropriate in certain circumstances **depending on the function for which such tool is used and other relevant factors.** In assessing whether disclosure is warranted, participants are encouraged to consider the extent to which (i) the output of an AI tool is to be relied upon in lieu of primary source material, (ii) the use of the AI tool could have a material impact on the proceeding, and (iii) the AI tool is used in a non-obvious and unexpected manner.

1-2. For instance, proactive disclosure may be warranted when (i) a party or an expert uses AI tools in the preparation of **evidentiary submissions, including expert testimony, witness testimony or documentary exhibits** and (ii) the use of such AI tools could have a **material impact** on the proceedings and/or their outcome.

1-3. Where disclosure is warranted, it should be **timely and sufficiently detailed to permit a reasoned objection or a request for further information.** Relevant details may include:

- 1) The name, version and relevant settings of the tool used;
 - 2) A short description of how the tool was used; and,
 - 3) In cases where reliance is placed on the output by a Generative AI tool, information regarding the complete prompt (including any template, additional context and conversation thread) and associated output.
-

[OPTION B]

1. Without limitation, disclosure may be appropriate **in the following circumstances:**

When a party or an expert (i) uses AI tools in the preparation of **submissions, expert opinions or other documents that are materially relied upon ([including evidence and demonstratives])** and (ii) the use of such AI tools could have **[an impact / a material impact]** on the proceedings and/or their outcome. In that case, they should include the following information in their disclosure:

- 1) The name of the tool used;
- 2) **Methodology** and a short description of how it was used (including, *e.g.*, prompts, instructions, or search terms);

*** Disclaimer: these Guidelines have been made publicly available in draft form for the purposes of receiving feedback and comments, and should not be used, adapted or relied on before the final version has been published by SVAMC.

<p>2. Should a party have reason to believe that another party or participant involved in the arbitration may have used AI tools in circumstances that warrant disclosure, it may submit an application explaining the reasons for such belief to the tribunal.</p> <hr/> <p>3. Arbitrators should make appropriate disclosure to the parties prior to using any AI tool in a manner that could be perceived as delegating any part of their decision-making function.</p> <hr/>	<p>2-1. Should a party have reason to believe that another party or an expert (i) used AI tools to prepare submissions, expert opinions or other documents that are materially relied upon ([including evidence and demonstratives]), and (ii) that the use of such AI tools could have [an impact / a material impact] on the proceedings and/or their outcome, it may submit an application explaining the reasons for such belief to the tribunal and request the disclosure of the information in Guideline 3.1.</p> <p>2-2. If a tribunal believes (i) that a party or an expert used AI tools to prepare their submissions, expert opinions or other materially relied-upon documents ([including evidence and demonstratives]) and (ii) that the use of such AI tools could have [an impact / a material impact] on the proceedings and/or their outcome, the tribunal may request the parties to disclose the information in Guideline 3.1, as well as other information they deem necessary.</p> <hr/> <p>3-1. If an arbitrator has used, is using, or intends to use AI tools, and deems its disclosure necessary, the arbitrator should include the information in Guideline 3.1 as well as other information they deem necessary.</p> <p>3-2. If an arbitrator has used, is using, or intends to use AI tools, in a way that could be perceived as delegating any part of their decision-making function, the arbitrator [should consider disclosing it] / [should disclose it] and provide the information in Guideline 3.1 as well as other information they deem necessary.</p> <hr/>
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

*** Disclaimer: these Guidelines have been made publicly available in draft form for the purposes of receiving feedback and comments, and should not be used, adapted or relied on before the final version has been published by SVAMC.

<p>4-1. The parties and the arbitral tribunal, should, at an early stage of the arbitration, consider the extent to which proactive disclosure of the use of AI tools should be required of parties, experts and the tribunal during the course of the proceedings. Any directives to this effect are without prejudice to a tribunal’s power to order disclosure related to the use of AI tools upon the request of a party or on its own motion.</p> <p>4-2. Decisions regarding disclosure of the use of AI tools shall be made on a case-by-case basis, considering, where applicable, the principles of transparency, due process, work-product privilege, and confidentiality of deliberative material.</p> <p>4-3. Counsel and other professionals retained by the parties should consider the extent to which proactive disclosure of their use of AI tools should be made to their clients.</p>	<p>4. Decisions regarding disclosure of the use of AI tools shall be made on a case-by-case basis, considering, where applicable, the principles of transparency, due process, work-product privilege, and confidentiality of deliberative material.</p>
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Commentary

Guideline 3 does not create any presumption in favour or against disclosure of the use of AI tools. Courts in certain jurisdictions have required parties and their attorneys to affirmatively disclose the use of Generative AI tools in preparing submissions, and/or certify the accuracy of submissions prepared using Generative AI. This Guideline does not impose mandatory disclosure or certification obligations by default. Some uses of AI by parties, experts and arbitrators may be uncontroversial and would not ordinarily warrant disclosure (*see* Examples of compliant and non-compliant uses of AI in arbitrations). To the extent that the arbitral tribunal, the parties, or the administering institutions consider it advisable, they may require proactive disclosure and/or certification in connection with the use of any Generative AI tool. As technology evolves and Generative AI tools become more accurate, however, and depending on the type of tool used, the need for such disclosures or certifications may need periodic reassessment.

Guideline 3 does recognise, however, that there are certain circumstances where disclosing the use of AI tools may be warranted to preserve the integrity of the proceedings or the evidence, although there are differences in the formulations proposed under Option A and Option B, respectively.

Option A identifies a range of factors that may be relevant in the assessment of whether disclosure is warranted, specifically whether (i) the output of an AI tool is to be relied upon in lieu of primary source

**** Disclaimer: these Guidelines have been made publicly available in draft form for the purposes of receiving feedback and comments, and should not be used, adapted or relied on before the final version has been published by SVAMC.*

material, (ii) the use of the AI tool could have a material impact on the proceeding, and (iii) the AI tool is used in a non-obvious and unexpected manner. Option A provides an example of when disclosure may be appropriate bearing these factors in mind but stops short of making it a requirement.

Option B goes a step further in proposing a two-prong test, providing for disclosure (i) when the output of AI tools is used to prepare or create materially relied-upon documents (including evidence, demonstratives, witness statements and expert reports) and (ii) when the output of that AI tool can have a material impact on the proceedings or their outcome. Disclosure in these cases should be proactive, at the party, expert or arbitrator's own initiative, but it can also be requested by a party by submitting an application to the tribunal.

When a party seeks disclosure on the use of AI tools from another, the materiality requirement seeks to discourage frivolous applications from disclosing fairly innocuous and uncontroversial uses of AI. Accordingly, under both Option A and B, a party seeking disclosure from another party should explain both (i) why it believes that an AI tool was actually relied upon in the proceedings and (ii) how it would materially impact the proceedings and/or their outcome.

Orders mandating disclosure of the use of AI tools and other related information may sometimes risk violating work-product privilege, which allows counsel to withhold strategies and materials prepared in connection with an arbitration. Thus, orders requiring AI disclosure must be approached carefully to prevent infringing work-product privilege, balancing considerations of transparency and due process with the need to preserve privilege or confidentiality.

Arbitrators should affirmatively disclose [Option A] or at least consider disclosing [Option B] the use of AI tools when that use could create the impression that an arbitrator is delegating part or all of their decision-making function (*see* Guideline 6 - Non-delegation of decision-making responsibilities) or, in any other situation the arbitrator deems necessary.

This Guideline aims to allow space for participants to explore and adopt AI tools without undue interference or retaliation, while maintaining checks and balances to safeguard the integrity of the arbitration process. The decision for disclosure should always be determined on a case-by-case basis, without reference to sweeping generalisations on the use of specific AI tools.

Ultimately, it is up to the parties and/or tribunal to specify the level of disclosure they want to institute for the proceedings, ideally at the outset of the arbitration, as expressly proposed in Option A.

The Guideline does not seek to regulate disclosure vis-a-vis clients as professional rules on the matter vary greatly by jurisdiction. However Option A explicitly encourages counsel and other professional service providers to consider whether their use of AI tools should be disclosed to their clients.

*** Disclaimer: these Guidelines have been made publicly available in draft form for the purposes of receiving feedback and comments, and should not be used, adapted or relied on before the final version has been published by SVAMC.

CHAPTER 2: GUIDELINES FOR PARTIES AND PARTY REPRESENTATIVES

GUIDELINE 4

Duty of competence or diligence in the use of AI

Party representatives shall observe any applicable ethical rules or professional standards of competent or diligent representation when using AI tools in the context of an arbitration.

Parties shall review the output of any AI tool used to prepare submissions to make sure it is accurate, from a factual and legal standpoint, as required by any applicable ethical rules or standards of competent representation (*see* Non-derogation of any mandatory rules). Parties and party representatives on record shall be deemed responsible for any uncorrected errors or inaccuracies in any output produced by an AI tool they use in an arbitration.

Commentary

Scope

This Guideline draws attention to some of the risks that may arise when party representatives delegate legal tasks (such as summarising cases, writing portions of briefs or oral submissions, or conducting legal research) to AI tools without reviewing the AI tool's output to make sure it is accurate, from a factual and legal standpoint.

As established in the Commentary to Guideline 1, certain Generative AI tools may be prone to errors and hallucinations, and their output can include inaccurate legal citations or mistakes in the presentation or interpretation of facts, evidence and legal authorities. Accordingly, this Guideline reminds party representatives (and particularly legal professionals) of their ethical and professional duty to review any work product created by, or with the help of, AI and remain responsible for inaccurate submissions made during an arbitration.

Guideline 4 does not impose an independent standard of review of party representatives' conduct. Rather, it contains *renvoi* to any applicable rules of professional conduct or responsibility to determine the level of diligence and reasonableness required when using AI tools. Party representatives on record will ultimately be deemed responsible for any non-compliance with this Guideline.

Consequences of non-compliance

Not all AI-induced errors are created equal. In some cases, an AI-induced error may be legitimately inadvertent, even after a reasonable review, or may be inconsequential or have no significant impact on

**** Disclaimer: these Guidelines have been made publicly available in draft form for the purposes of receiving feedback and comments, and should not be used, adapted or relied on before the final version has been published by SVAMC.*

the arbitration. In other cases, AI-induced errors and hallucinations can compromise the integrity of the proceedings, or result in a skewed presentation of the facts, the law or the evidence (*see* Guideline 5 - Respect for the integrity of the proceedings and the evidence).

The tribunal can take these factors into account when deciding how to address submissions containing AI-induced errors and inaccuracies.

*** Disclaimer: these Guidelines have been made publicly available in draft form for the purposes of receiving feedback and comments, and should not be used, adapted or relied on before the final version has been published by SVAMC.

GUIDELINE 5

Respect for the integrity of the proceedings and the evidence

Parties, party representatives and experts shall not use any forms of AI in ways that affect the integrity of the arbitration or otherwise disrupt the conduct of the proceedings.

Parties, party representatives and experts shall not use any form of AI to falsify evidence, compromise the authenticity of evidence, or otherwise mislead the arbitral tribunal and/or opposing party(ies).

Commentary

This Guideline prohibits any use of AI that compromises the integrity of the arbitration or the authenticity of evidence. While deploying AI can enhance the efficiency of arbitration proceedings, its potential misuse may disrupt due process and corrupt an arbitral tribunal's findings.

The duty to protect the integrity of the proceedings and not to submit false or adulterated evidence already exists in arbitration. Fraudulent behaviour and misconduct, such as submitting false documents or resorting to so-called "guerilla tactics", can occur with or without the use of AI.

Advancements in AI, however, particularly in Generative AI and deep fakes, can heighten the risks of manipulated or false evidence, making it significantly easier to create fake evidence that can appear strikingly convincing to the naked eye or which can sometimes be virtually indistinguishable from authentic versions. It can also make it more costly or difficult to detect any such manipulation through forensic and other means.

This Guideline reminds parties to be aware and vigilant of these heightened risks while emphasising the importance of ensuring the fairness and integrity of the proceedings when using AI. Parties, party representatives and experts should simply not use AI tools to fabricate evidence, distort evidence, or compromise the integrity of the proceedings under any circumstances.

If the arbitral tribunal determines that a party has violated this Guideline, it may consider, in addition to any other measures available under the applicable arbitration rules or the *lex arbitri* (such as, for example, striking the evidence from the record, or deeming it inadmissible), taking the infringing party representatives' conduct into account in its assignment of the costs of the arbitration.

*** Disclaimer: these Guidelines have been made publicly available in draft form for the purposes of receiving feedback and comments, and should not be used, adapted or relied on before the final version has been published by SVAMC.

CHAPTER 3: GUIDELINES FOR ARBITRATORS

GUIDELINE 6

Non-delegation of decision-making responsibilities

An arbitrator shall not delegate any part of their² personal mandate to any AI tool. This principle shall particularly apply to the arbitrator’s decision-making function.

Commentary

Non-delegation of personal mandate

This Guideline underlines the critical principle that an arbitrator’s mandate, especially their ultimate decision-making function, is personal and non-delegable. This Guideline does not forbid the use of AI tools by arbitrators as an aid to discharge their duty to personally analyse the facts, arguments, evidence and the law and issue a reasoned decision.

While AI tools can assist in managing information, analysing data, and predicting outcomes, they should not replace the human judgement, discretion, responsibility, and accountability inherent in an arbitrator’s role. Therefore, arbitrators must be mindful that they are not inadvertently delegating part of this personal mandate to the AI tool.

Under this Guideline, arbitrators need to review the output produced by any AI tool to ensure it is accurate and shall take responsibility for any errors or inaccuracies. If an arbitrator uses a Generative AI tool to assist in the analysis of the arguments or the drafting of a decision or award, the arbitrator cannot simply reproduce the AI’s output without making sure it adequately reflects the arbitrator’s personal and independent analysis of the issues and evidence at hand.

This Guideline reminds arbitrators that, even as technology evolves, their personal responsibility in rendering decisions remains paramount. AI can enhance efficiency and provide insights, but the arbitrator must make the ultimate decisions, preserving the human element essential to the fairness and integrity of arbitration proceedings. At all times, the arbitrators remain responsible for the use of AI during the arbitration.

² The terms “their”, “they”, and “them” as used in these Guidelines in relation to any of the individual participants in an arbitration are used as singular, gender-neutral pronouns.

*** Disclaimer: these Guidelines have been made publicly available in draft form for the purposes of receiving feedback and comments, and should not be used, adapted or relied on before the final version has been published by SVAMC.

GUIDELINE 7

Respect for due process

An arbitrator shall not rely on AI-generated information outside the record³ without making appropriate disclosure to the parties and, as far as practical, allowing the parties to comment on it.

Where an AI tool cannot cite sources that can be independently verified, an arbitrator shall not assume that such sources exist or are characterised accurately by the AI tool.

Commentary

This Guideline focuses on the principle of due process in using AI in arbitration. It emphasises the arbitrator’s duty to disclose any reliance on AI-generated outputs outside the record that influence their understanding of the case, to the extent that any outputs are used, allowing parties the opportunity to comment. This approach ensures transparency and upholds the parties’ right to be heard.

At the same time, it acknowledges that disclosure requirements may vary depending on the specific AI application used.

The Guideline also stresses the arbitrator’s responsibility to avoid assuming the existence of authoritative sources from AI outputs. It prompts arbitrators to evaluate the reliability of AI-derived information independently and critically.

³ Some civil law jurisdictions recognise the principle of *iura novit arbiter*, or the “arbitrator knows the law”, pursuant to which arbitrators may have the authority to apply laws, case law and precedents not cited by the parties. This principle has also been applied in investment treaty cases and by the International Court of Justice. The extent of this authority may vary depending on the jurisdiction. However, this Guideline does not preclude in any way the application of the principle of *iura novit arbiter*, where appropriate.

*** Disclaimer: these Guidelines have been made publicly available in draft form for the purposes of receiving feedback and comments, and should not be used, adapted or relied on before the final version has been published by SVAMC.

EXAMPLES OF COMPLIANT AND NON-COMPLIANT USES OF AI IN ARBITRATIONS

For each Guideline, this section offers a few practical examples of both compliant and non-compliant uses of AI in international arbitration.

These instances are not exhaustive but illustrative, encouraging thoughtful use of AI while ensuring the principles of fairness, integrity, and equality are preserved in arbitration proceedings. Ultimately, whether the use of AI in international arbitration in a given case is appropriate or not will need to be determined on a case-by-case basis.

GUIDELINE 1 - Understanding the uses, limitations and risks of AI applications	
Compliant	Non-compliant
Using a specialised AI tool to conduct research on potential arbitrators or experts for a case, being mindful of the AI tools' limitations and evaluating the results accordingly.	Using AI tools to select arbitrators or experts for a case without human input and without assessing the AI tool's selection critically and independently or controlling for biases and other limitations.
GUIDELINE 2 - Safeguarding confidentiality	
Compliant	Non-compliant
Using AI tools for routine non-confidential tasks, such as meeting scheduling.	Submitting confidential information to a third-party AI tool without proper authorisation and where the terms of use for such tool allow logging of inputs/outputs and sharing them with third parties.
Using AI tools to research or summarise legal authorities in a third-party database, provided there is no sharing of confidential information.	
GUIDELINE 3 - Disclosure and protection of records	
Compliant	Non-compliant

*** Disclaimer: these Guidelines have been made publicly available in draft form for the purposes of receiving feedback and comments, and should not be used, adapted or relied on before the final version has been published by SVAMC.

Using AI tools to generate document summaries for internal use, creating indices or proofreading drafts without disclosing it.	Using an AI tool to calculate damages without disclosing it and providing information to critically assess reliance on the AI tool's output.
Using AI tools to identify and select the documents potentially relevant and responsive to document production requests while disclosing the manner in which such tool was used in a way that would permit the opposing party to make an informed objection.	As an arbitrator, use an AI tool to “score” or otherwise compare the persuasiveness of parties’ submissions without disclosing it (assuming the arbitrator has checked the accuracy of the AI tool’s output).
GUIDELINE 4 - Duty of competence or diligence in the use of AI (parties)	
Compliant	Non-compliant
Using AI tools to assist with drafting language for pleadings or written submissions where the final work product is fully source-checked and vetted for accuracy from a factual and legal standpoint.	Using AI tools to draft pleadings or written submissions without checking the accuracy of their output from a factual and legal standpoint.
Using specialised AI tools to find or summarise relevant cases, vetting the accuracy of the descriptions before incorporating them into pleadings.	Using AI tools to summarise cases and “copy-paste” them into pleadings without verifying whether the AI’s output may contain any errors.
Using AI tools to assist in the preparation of cross-examination questions or find inconsistencies in witness statements.	
GUIDELINE 5 - Respect for the integrity of the proceedings and the evidence (parties)	
Compliant	Non-compliant

**** Disclaimer: these Guidelines have been made publicly available in draft form for the purposes of receiving feedback and comments, and should not be used, adapted or relied on before the final version has been published by SVAMC.*

<p>Using AI tools to produce demonstratives, such as 3D and other graphic representations, where the demonstratives are based on evidence in the record and the accuracy of the representation can be challenged by the opposing party by accessing the referenced source data.</p>	<p>Using AI tools to falsify or otherwise manipulate documents submitted as evidence.</p>
<p>GUIDELINE 6 - Non-delegation of decision-making responsibilities (arbitrators)</p>	
<p>Compliant</p>	<p>Non-compliant</p>
<p>As an arbitrator, using an AI tool capable of providing accurate summaries and citations to create a first draft of the procedural history of a case, or generate timelines of key facts, and then double-checking accuracy of the AI tools' output with underlying sources and making other appropriate edits.</p>	<p>As an arbitrator, using an AI tool to provide an assessment of the parties' submissions of evidence and incorporate such output into a decision without conducting an independent analysis of the facts, the law and the evidence to make sure it reflects the arbitrator's personal and independent judgement.</p>
<p>GUIDELINE 7 - Respect for due process (arbitrators)</p>	
<p>Compliant</p>	<p>Non-compliant</p>
<p>As an arbitrator, using AI tools to distil or simplify technical concepts to come up with technically accurate or relevant questions for the parties or experts during the hearing.</p>	<p>As an arbitrator, using Generative AI tools to conduct independent research into the substance of the dispute and base their decision on such generated outputs without disclosing it to the parties and providing them an opportunity to comment.</p>

**** Disclaimer: these Guidelines have been made publicly available in draft form for the purposes of receiving feedback and comments, and should not be used, adapted or relied on before the final version has been published by SVAMC.*

MODEL CLAUSE FOR INCLUSION IN PROCEDURAL ORDERS

The Tribunal and the parties agree that the Silicon Valley Arbitration and Mediation Center’s Guidelines on the Use of Artificial Intelligence in International Arbitration (**SVAMC AI Guidelines**) shall apply as a reference framework to all participants in this arbitration proceeding.