

Kluwer Arbitration Blog

From Draft to Published Version: A Commentary on the Changes Implemented in the SVAMC AI Guidelines

Felix Kasseroler · Friday, June 21st, 2024

The Silicon Valley Arbitration and Mediation Center's (SVAMC) Guidelines on the Use of Artificial Intelligence in Arbitration ("Guidelines") strive to become the first set of rules governing artificial intelligence ("AI") recognized within the international arbitration community. These have previously been briefly discussed [here](#), [here](#) and [here](#). After more than a year of preparation, including a six-month comment period, the Guidelines were published on 30 April 2024 and have been tailored to the needs of the international arbitration community. This includes a fairly broad and future-oriented definition of AI, provisions on safeguarding confidentiality and furthermore, ensuring the integrity of the arbitral proceeding.

However, before having a deeper look into the Guidelines themselves, one question needs to be answered: is there even a need for further rules? The answer is rather self-explanatory: with AI tools becoming more accessible to the general public over the last couple of years, the number of uses and users has grown exponentially. However, this growth also harbors risks and limitations. One problem is where not even the creators of those tools are able to fully comprehend the decisions made by their own AI tool (which is known as the "black-box problem"). Another arises from how generative AI models leverage intricate probabilistic computations to produce outputs that exhibit a remarkable degree of authenticity, even though the underlying mechanism is more akin to a sophisticated statistical estimation than a transparent reasoning process. This thus gives rise to "hallucinations", which occur where generative AI models lack information to produce a coherent and fluent output. Instead of highlighting this lack of information to the user, these models produce an output based on mathematical probabilities, without assessing its accuracy.

Therefore, these tools should, in line with Guideline 1 of the AI Guidelines, be used with the utmost care and furthermore, in accordance with applicable jurisdictional and/or international standards. The Guidelines aim to govern these technological developments and try to act as a general rulebook regarding the implementation and use of AI in international arbitral proceedings. In spite of specific wording such as "general principles", the Guidelines – as stated in the preliminary provisions – do not derogate from any mandatory rules but seek rather to guide and "accommodate case-specific circumstances". Where rules and regulations (e.g. laws, treaties, domestic statutes, ethical rules, etc.) on the use of AI in international arbitration apply, whether on a domestic or international level, the Guidelines do not detract or derogate from any applicable mandatory rules and regulations but rather act as a supplementary international standard.

This article does not aim to go over the recently published Guidelines in their entirety. Instead, it

tries to highlight the most significant changes and additions to the draft that has been public since 31 August 2023.

Preliminary Matter: The Definition of AI

Due to the extensive and multi-purpose uses and the lack of a single, internationally recognized definition, the Guidelines set out a rather broad definition of what counts as AI. It aims to include “existing and future foreseeable types of AI” but is phrased narrowly so as not to cover every computer-assisted tool such as spellcheck programs. It therefore encompasses the publicly known generative AI as well as evaluative or discriminative AI. However, it should once again be noted that the broad definition might also include future AI tools which therefore would have to be evaluated on a case-by-case basis.

The General Obligation to Disclose AI Use

First and foremost, the general obligation to disclose the use of AI, found in Guideline 3, has been extensively discussed within the international arbitration community (e.g. during the Paris Arbitration Week 2024 panel discussions: “Exploring the Use of Artificial Intelligence in Arbitral Award Writing” on 20 March 2024 and “Using AI in Arbitration Proceedings: An Update on the SVAMC AI Guidelines” on 22 March 2024).

The final text of Guideline 3 reads as follows:

“Disclosure that AI tools were used in connection with an arbitration is not necessary as a general matter.

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.

Where appropriate, the following details may help reproduce and evaluate the output of an AI tool:

- 1. the name, version, and relevant settings of the tool used;*
- 2. a short description of how the tool was used; and*
- 3. the complete prompt (including any template, additional context, and conversation thread) and associated output.”*

In relation to this Guideline, the draft of 31 August 2023 contained two different versions of an obligation to disclose the use of AI in arbitral proceedings (previously discussed [here](#)). However, some within the arbitral community did not agree with the proposed content. This included several audience members of the two above-mentioned panel discussions during the Paris Arbitration Week 2024. In particular, some from larger law firms that were already using AI and legal tech in their daily work tended to disagree with the obligation to disclose.

Due to the adverse reaction to Guideline 3, the SVAMC decided in the final version to make disclosure subject to a case-by-case decision rather than an obligation that is generally applicable. The commentary on Guideline 3 states for example that disclosure could mitigate “concerning uses of AI that would otherwise fall under Guideline 5”. If, however, disclosure is proclaimed necessary in a specific case, the provision also provides the information needed for full disclosure such as name, version and relevant setting of the tool used, a short description of the specific use and the complete prompt.

Safeguarding Confidentiality

In the final version, the SVAMC also added the obligation to “redact or anonymize materials submitted to an AI tool” in Guideline 2 where appropriate. While the draft provision already guided users to pay special attention to an AI tool’s “policies on recording, storage and use of prompt or output histories and of any other confidential data submitted to the AI tool”, the obligation to redact or anonymize where appropriate marks a higher threshold towards disclosing sensitive information to an AI tool regardless of specific confidentiality policies. The interpretation of “[w]here appropriate” in this specific context will have to be determined in practice in order to successfully apply this provision.

Arbitrator-specific Concerns

The final text of Guideline 6 provides:

“An arbitrator shall not delegate any part of their personal mandate to any AI tool. This principle shall particularly apply to the arbitrator’s decision-making process. The use of AI tools by arbitrators shall not replace their independent analysis of the facts, the law, and the evidence.”

With regard to the provisions specific to arbitrators, Guideline 6 was supplemented with the underlined words to highlight the duty of an arbitrator to independently analyze the facts, the law, and the evidence of a case. The inclusion in the Guideline itself rather than in the commentary as per the draft, shows the importance of the arbitrator’s obligation to draft awards. This does not in any way preclude an arbitrator from generally consulting AI. However, the commentary emphasizes the importance of an arbitrator’s obligation to not delegate parts of the arbitrator’s personal mandate which includes rendering a decision based on “human judgement, discretion, responsibility and accountability”. It obligates the arbitrator to ensure that any output made by AI is accurate and valid, as if it were themselves who made it. In regard to this, the commentary also states that an arbitral tribunal must personally review any output or draft a decision and will not be released from this duty.

Lastly, while Guideline 7 remains the same, it now contains additional commentary on the reliability of any output by an AI tool. For context, Guideline 7 prevents an arbitrator from utilizing AI-generated output that originates outside of the formal procedural record, unless prior to any such use, the arbitrator provides the parties with adequate disclosures regarding the nature and

source of the information and an opportunity to comment on such use. In the final version, the commentary also states that arbitrators, like parties and party representatives, need to independently and critically assess any output derived from an AI tool.

Conclusion

The SVAMC Guidelines mark the first approach taken to the governance of AI use in international arbitral proceedings. Time will tell how the Guidelines will be received by the international arbitration community, whose adoption of the same may be made slightly easier by the model clause contained in the Guidelines to facilitate their inclusion in procedural orders.

To make sure you do not miss out on regular updates from the Kluwer Arbitration Blog, please subscribe [here](#). To submit a proposal for a blog post, please consult our [Editorial Guidelines](#).

2024 Summits on Commercial Dispute Resolution in China

17 June – Madrid
20 June – Geneva

Register Now →



This entry was posted on Friday, June 21st, 2024 at 8:15 am and is filed under [Artificial Intelligence](#), [SVAMC AI Guidelines](#)

You can follow any responses to this entry through the [Comments \(RSS\)](#) feed. You can leave a response, or [trackback](#) from your own site.

