



Navigating Legal And Ethical Conundrums of Using AI-Generated Content (AI-GC) Systems In Arbitration

Mohamad Fateh Labanieh¹, Mohammad Azam Hussain², Zainal Amin Ayub³, Hussein Alaa Al-Azzawi⁴

¹ Universiti Utara Malaysia, Malaysia

² Universiti Utara Malaysia, Malaysia

³Universiti Utara Malaysia, Malaysia

⁴Universiti Utara Malaysia, Malaysia

m.fateh.labanieh@uum.edu.my

Abstract. Developments in Artificial intelligence (hereinafter referred to as “AI”) technologies and the prompt transformation brought by such technologies lead to several critical questions and dilemmas. By using doctrinal legal research methodology and employing critical and analytical techniques, this article aimed to examine the legal and ethical conundrums of using AI-generated content (hereinafter referred to as “AI-GC”) systems in arbitration. It found that although AI systems, including AI-GC systems, would play a vital role in arbitration and the legal industry, there are several legal and ethical conundrums in using those systems in arbitration. Therefore several recommendations and insights have been provided to enlighten not only the users, such as arbitrators but also the Malaysian lawmakers. Considering them will bring added value to the reputation of Malaysia as an advanced jurisdiction and significantly enhance the greater good of using AI technologies in the arbitration industry.

Keywords: Arbitration; Dispute Resolution; Artificial intelligence (AI)

1 Introduction

Arbitration is the preferred method for resolving cross-border disputes on several grounds. These include expediency, competence, impartiality and enforceability. Furthermore, arbitration differs from litigation due to its confidentiality. To illustrate more, in contrast to litigation, this confidentiality extends to hearings and parties’ submissions.

Due to the accelerated development of AI technologies, the legal industry, including arbitration, is undergoing profound changes [1]. There is no singular definition of AI that is universally accepted. This is mainly because AI is a term that encompasses a variety of subjects, including machine learning (ML), natural language processing (NLP), and others. Therefore, Schuett (2019) recommended that policymakers “should not use the term “artificial intelligence” for regulatory purposes because there is no definition of AI which meets the requirements for legal definitions. Instead, they should

define certain designs, use cases or capabilities following a risk-based approach” [2] page1).

According to John McCarthy, AI is “the science and engineering of making intelligent machines”, ‘intelligence’ being ‘the computational part of the ability to achieve goals in the world” [3]. Moreover, Calo (2017) argued that “AI is best understood as a set of techniques aimed at approximating some aspect of human or animal cognition using machines” [4]. Additionally, the World Intellectual Property Organisation (WIPO) asserts that;

“AI is generally considered to be a discipline of computer science that is aimed at developing machines and systems that can carry out tasks considered to require human intelligence. Machine learning and deep learning are two subsets of AI. In recent years, with the development of new neural networks techniques and hardware, AI is usually perceived as a synonym for “deep supervised machine learning”.

AI-GC is a type of AI system that can generate, produce and create content, such as images, text, and videos. Specifically, it refers to the process of teaching a machine learning model (hereinafter referred to as “MLM”) with vast amounts of text data to create novel and original compositions from pre-existing content and large volumes of text and data scraped primarily from the internet repositories [5]. The AI-GC system will systematically learn from the internet repositories data that is fed to develop its communication, artistic, and textual expression skills and then uses it (“what it knows”) to produce fresh outputs based on written prompts from users.

Several AI-GC systems are designated to carry out several tasks. For instance, ChatGPT, Google Bard, ChatSonic, and Jasper AI (AI-generated content), DALL-E and AI Stable Diffusion (AI-generated images), Amper Music and AIVA (AI-generated music), AlphaCode (AI-generated code). These systems have increased noticeably across various industries because of their massive capacity to revolutionise traditional content review and creation methods, opening the door to enhance efficiency and productivity [6]. From the legal point of view, there are several cases where national judges in different jurisdictions, such as Colombia [7], Pakistan [8], India [9], Bolivia [9], have used AI-GC systems, specifically, ChatGPT to make a court ruling, justify their decisions and aid them in reaching decisions.

Using AI systems in arbitration is not expressly prohibited by the Malaysian Arbitration Act 2005 (Act 646) (hereinafter referred to as “Act 646”) as a result of applying the principle of party autonomy [10]. In short, an essential advantage in arbitration is party autonomy. In accordance with this principle, the parties to the arbitration are free to decide on many aspects of the arbitral procedure, subject to the limitations given by the mandatory rules of the relevant procedural law “lex arbitri”. However, it is necessary to know that AI-GC systems, such as ChatGPT, have remarkable drawbacks and raise essential conundrums.

Therefore, to ensure fairness in arbitration, proper consideration should be given to legal and ethical conundrums associated with using AI-GC systems in arbitration. This article employed doctrinal legal research methodology to examine the legal and ethical conundrums of using AI-GC systems in arbitration. It highlighted the positive role of using AI systems, including AI-GC systems, in the legal industry. Then it examines the legal and ethical conundrums of using those systems in arbitration.

Two overarching objectives essentially drive this endeavour. First, they provide arbitrators with an in-depth understanding of the potential adverse effects of using AI-GC systems in arbitration. Second, informing Malaysian legislators about the legal issues posed by using AI-GC systems in arbitration encourages them to take action to keep pace with the rapid breakthroughs in the AI industry.

2 Methodology

This article is based on doctrinal legal research methodology. Primary data and secondary data are used and analysed using critical and analytical approach. It is hereby submitted that this article approach is adequate for meeting the purpose of the study.

3 The Positive Role of Using AI Systems, Including AI-GC Systems, in the Legal Industry

Before analysing the legal and ethical conundrums of using AI-GC systems in arbitration, it is important to briefly highlight the positive role of using AI systems, including AI-GC, in the legal industry. The following discussion is primarily concerned with the applicability of current AI technology in the legal industry, including the arbitration industry (if any).

By increasing efficiency, lowering costs, and improving accuracy, AI systems, such as AI-GC systems, have the potential to transform the arbitration industry completely. This is because of the automation power of AI-GC systems. One of the AI-GC systems is known as Cecilia. This system's design permits attorneys to ask inquiries in natural language and receive responses with specific citations to supporting evidence from the private DISCO eDiscovery databases [11]. AI-GC systems, such as ChatGPT, can be used to analyse, summarise [12], and write an arbitral award. Also, Kira System is an ML software that extracts, identifies and analyses content in contracts and documents [13].

In addition, AI-GC systems, such as ChatGPT, also have other significant benefits, most notably improved consistency. Specifically, contrary to humans, who may exhibit performance variability, AI-GC systems are able to complete the tasks assigned regardless of the situation, time of day, and unnecessary factors, such as the food break. To illustrate more, a study uncovered an essential pattern among more than one thousand (1000) judicial decisions. The study demonstrated that many applications are consistently denied on average. Following the judges' daily meal break, an unusually high number of favourable decisions were observed to be rendered [14]. This is one of the few reasons AI-GC systems are considered more effective and secure from the unnecessary factors affecting the tasks assigned to those systems, compared with humans. Furthermore, RavelLaw and ArbiLex are AI systems capable of predicting potential and future outcomes, such as arbitral awards. With the aid of predictive justice, measuring uncertainties and increasing the likelihood of favourable outcomes in arbitration

disputes becomes easier. AI systems can also help select arbitrators, such as Arbitrator Intelligence and Global Arbitration Review Arbitrator Research Tool. Moreover, AI systems may be successful in cross-examination [15] and translation as applied by Hong Kong's Electronic Business-Related Arbitration and Mediation Platform (eBRAM) [16]. In light of the aforementioned, it is worth noting that even though AI systems, including AI-GC systems, have an opportunity to improve productivity and efficiency, as well as decrease costs and time wasting, the reliability and integrity of the arbitration procedures and the interests of the parties involved, should always be maintained.

4 The Legal Conundrums of Using AI-GC Systems in Arbitration

This section examines the legal issues underlying the use of AI-GC in arbitration. It focuses mainly on two important aspects: first, the inherent potential for AI-GC systems to generate biased outcomes and content, and second, the examination of copyright issues associated with using such systems.

4.1 The Potential for Providing Biased Outcomes and Contents

The arbitrator in arbitration must be impartial and independent (free from bias) (section 13 (8) (b) of Act 646); otherwise, a fair resolution will not be achieved [17]. Furthermore, if there is any justifiable doubt regarding the arbitrator's independence or impartiality, any of the parties to arbitration has the right to challenge that arbitrator (section 14 (3) (a) of Act 646). The procedures to challenge the biased arbitrator are mentioned in section 15 of Act 646.

Regardless of the previous facts, arbitration is not protected from the risk of arbitrator's bias [18], such as ethnic or cultural bias [19], because, on some occasions, the arbitrator might not impartially perform his/her obligations and tasks. Moreover, the bias is an unnoticeable phenomenon [20], and it can be either apparent or actual. The difference between the "actual bias" and "apparent bias" is that the former occurs where there is a situation proving that the arbitrator has been actually biased in reaching his/her arbitral award. However, apparent bias occurs when there is reasonable doubt or apprehension that the arbitrator might appear biased. For example, the mere fact that the arbitrator has a personal interest in the dispute's outcome is enough to challenge such an arbitrator [21]. Further, it is difficult to prove the actual bias because the arbitrator will not explicitly show his/her favouritism and preference towards one of the parties [18]. Therefore, a party, who wants to challenge the arbitrator, will usually strive to prove the arbitrator's apparent bias [22]. The test for bias differs from one country to another. For instance, in Malaysia, the test of bias is based on the "real danger of bias" (*MPPP v. Syarikat Bekerjasama sama Serbaguna Sungai Gelugor* [1999] 3 MLJ 1; *Dato' Tan Heng Chew v. Tan Kim Hor & Anor* [2006] 1 CLJ 577), unlike English law that adopts the test of "real possibility of bias" (*Sierra Fishing Co & Others v. Farran & Others* [2015] EWHC 140).

In the context of this article, the issue of bias in AI has become a developing concern. The term bias means the unequal or unfair treatment of certain groups of individuals based on their ethnicity, gender, or other characteristics. Bias in AI can have severe repercussions, as it can worsen the unfair treatment of marginalised individuals. Due to the deeply entrenched of biased that is used to train the algorithms, AI systems, particularly AI-GC systems, also have the potential to deliver decisions, outcomes, and content that are systematically unfair and biased in favour of some people. This poses a substantial problem for arbitrators who are bound to identify possible biases and embark on actions to verify the precision and neutrality of the answers provided by AI-GC systems, such as ChatGPT. To elaborate further, AI-GC systems, such as DALL-E 2 and Stable Diffusion, encountered challenges in generating images of “older couples of colour”, specifically until the word “poor” was incorporated [23].

Furthermore, an essential study called “gender shades”, disseminated in 2018, found that popular facial recognition systems most accurately detected males with lighter skin and had the highest errors detecting females with darker skin. In the following year, another groundbreaking study revealed a shocking finding. A clinical algorithm employed by many hospitals to prioritise patient care has demonstrated racial bias [24]. In addition, a risk assessment tool called Correctional Offender Management Profiling for Alternative Sanctions algorithm (COMPAS) is used in the United States of America to help judges decide whether to set bail for defendants by allegedly predicting the possibility that they will commit the same crime again. However, because the past information and data utilised to train the system contained established prejudices, COMPAS replicated and maintained similar biases in its predictions [25] [26].

From the authors’ personal experience, we just asked ChatGPT the following question; “ChatGPT, may you please give me a list of books on arbitration and their authors “regardless of race and language.” The following figure shows the answer provided by ChatGPT.

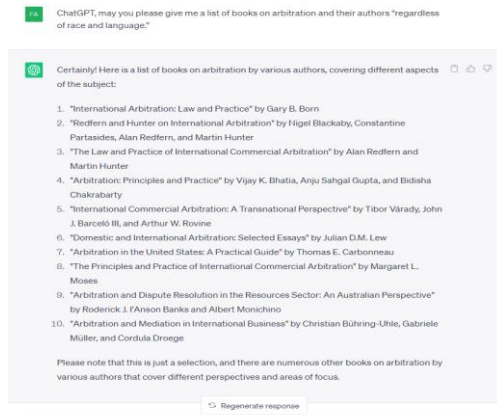


Fig. 1. Sourced from ChatGPT

Considering the foregoing figure, the result was frustrating because all of ChatGPT's recommendations had been written in English. There are no recommended readings from Muslim, Asian or Malaysian authors, such as Datuk Professor Sundra Rajoo. In addition, the majority of the recommended readings belong to white-man authors. Therefore, it is argued that using data from Western literature and books to train AI-GC systems, such as ChatGPT, proved the ethnic, regional, gender and religious biases in the output generated by ChatGPT.

Consequently, one may argue that the AI system lacks the capacity for any form of personal relationships, whether friendly or hostile [27]. However, it is argued that AI systems, such as AI-GC systems, can provide biased content. Therefore, there is an urgent necessity to address the issue of biases in the AI-GC systems appropriately. This can be achieved by the following. First, using various and diverse data when training the AI systems, such as the AI-GC system (data represent the whole population and balance). Second, testing the AI systems, such as AI-GC systems, for any potential bias before deploying them. Third, designing and developing the AI-GC systems in a way that considers the users' needs. Four, ensuring that the AI system, such as the AI-GC system, is explainable and transparent. Considering this point, the users (arbitrators) and programmers of AI-GC systems should understand how and why AI-GC systems generate such content (arbitral award) before using it. Understanding the fundamental causes of preferences can help in gaining essential knowledge of the intricate operations of AI systems and enhance the use of AI systems, including AI-GC systems, in the arbitration industry.

4.2 The Potential for Copyrights Issue

As AI-GC systems become increasingly integrated into our daily lives, copyright disputes are growing and becoming common. In a recently filed case, *Andersen v. Stability AI*, Civil Action No. 3:23-cv-00201 (N.D. Cal.), a group of artists sued Stability AI, Midjourney, and Deviant Art. They claim these AI companies had infringed their intellectual copyrights by unlawfully training their AI systems on "web-scraped images". Although this issue of copyright in using AI-GC systems is not directly related to arbitration, it is important to examine the Malaysian legal position specifically, whether the content or work produced entirely by AI-GC systems can be copyrighted or not.

After analysing the Copyright Act 1987 (Act 332) (hereinafter referred to as "Act 332"), it is witnessed that Act 332 is clear about refusing to grant copyright for the content generated entirely by AI-GC systems. Act 332 is implicitly presumed that the author of the copyright should be a human, "a natural person", because it uses human-associated pronouns, such as he/she or his/her, as indicated under section 17(1) of Act 332. It states that "Except as otherwise provided in this Act, copyright in any literary, musical or artistic work that subsists in such work under this Act shall subsist during the life of the author and shall continue to subsist until the expiry of a period of fifty years after his death." Another justification can be seen section 3 of the Act 332, where the terms "author" and "qualified person" are not extended to cover the non-human, such as AI-GC system.

Regarding whether the content or work produced with the assistance of AI-GC systems can be copyrighted or not. It was also witnessed that section 7 (3) of Act 332 highlights the criteria for copyright eligibility of a literary, musical, or artistic work. In order for a work to qualify for copyright protection, two conditions must be met. Firstly, sufficient effort must be put into making the work original in character. This means the work should possess a certain level of creativity or originality, significantly distinguishing it from existing works. It implies that a mere reproduction or duplication of an existing work may not meet the threshold for copyright protection. Secondly, the work must be fixed in a tangible form, such as being written down, recorded, or otherwise reduced to material form. This requirement emphasises the need for the work to exist in a concrete and perceptible medium. Once the work has been captured in a physical or digital format, it becomes eligible for copyright protection.

In view of the above, a critical question remained without an answer. The question is about if the user of the AI-GC system modified the output content generated by such system in a way that satisfies and complies with section 7 (3) of the Act 332, would that content be eligible for copyright? Therefore, lawmakers should adopt a well-defined stance on this issue.

5 The Ethical Conundrums of Using AI-GC Systems in Arbitration

This section examines the ethical issues underlying the use of AI-GC in arbitration. It focuses mainly on two essential aspects: first, the inherent potential for AI-GC systems to provide fake and incorrect content, and second, the lack of reasoning in AI-GC systems.

5.1 The Potential for Copyrights Issue

AI systems, including AI-GC systems, have the potential to improve tasks performed by humans, consequently improving cost-effectiveness and efficiency, as they have the capacity to generate text free of grammatical errors. Nevertheless, it is essential to know that the AI-GC systems used in such procedures are not flawless. Therefore, they might generate misinformation and fake outcomes but sound convincing to deceive humans. In light of the aforementioned, Galactica (similar to ChatGPT) is an AI large language model (hereinafter referred to as “LLM”) that is able to summarise articles, solve equations, and perform other scientifically-oriented actions. However, it was discovered that Galactica cited fictitious authors for non-existent articles [28]. Furthermore, BlenderBot, a chatbot, had unnatural conversations and made offensive and incorrect statements [29]. In addition, recently, The Guardian (a British daily newspaper) discovered that ChatGPT used and cited fake and not-existed articles in the Guardian database [30]. Another example can be seen when ChatGPT was asked to write an article, and the article was filled with false information [31].

Moreover, from Figure 1 (provided earlier), the authors also identified that ChatGPT had provided fake and incorrect information. Specifically, the recommended reading number 6 (“Domestic and International Arbitration: Selected Essays” by Julian D.M. Lew) is wrong, and the correct title should be (“Carbonneau on International Arbitration: Collected Essays” by Julian D.M. Lew).

In the context of this article, the arbitrator may use AI-GC systems to write an arbitral award or summarise and analyse the parties’ arguments and other documentation. This raises a concern about the reliability of the outcomes, such as an arbitral award generated by an AI-GC system, along with the accuracy level of the summarisation and analyses produced by such a system. With respect to this observation, it is argued that since AI systems, including AI-GC systems, can benefit those involved in arbitration, the arbitration industry should welcome AI adoption rather than dismiss it. However, it is advisable to exercise caution when the arbitrators decide to use AI-GC systems to assist them. This can be achieved by double-checking the results to ensure a high-quality resolution, greater certainty, and sustainability rather than exclusively relying on results or outcomes generated by AI-GC systems.

5.2 The Lack of Reasoning in AI-GC Systems

The primary function of the arbitral members is to decide the dispute before them by issuing an award [32]. Generally, the arbitral award contains several elements, among them the majority of the arbitrators’ signatures (only if the arbitral tribunal contains more than one arbitrator) and the reason for the absence of other signatures (section 33 (2) of Act 646). In fact, the main reason for the arbitrator’s signatures is to provide definitive proof that the arbitrator has rendered the award and to ensure that the arbitral award is eligible for enforcement.

In addition, the arbitral award should include the reasoning (justification and explanation) upon which it was based (section 33 (3) of Act 646)). The same is required from the international perspective. For instance, article 30(2) of the UNCITRAL Model Law on International Commercial Arbitration 2006 stated that “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”. It is necessary to emphasise that there exist numerous objectives in delivering a reasoned arbitral award, encompassing the possibility of the parties to understand how and why the award is rendered.

In fact, the outputs, such as arbitral awards, generated by AI systems, including AI-GC systems, are unexplainable. This is due to the fact that these systems are considered opaque and rarely “have any concrete sense of how or why a particular classification has been arrived at from inputs” [33]. Against this backdrop, consider deciding on a small claim arbitration dispute to illustrate this idea. If you asked an AI system why you decided the dispute in favour of A, not B, an AI system might say that the argument provided by A is stronger than the argument provided by B. However, suppose one digs deeper into this issue. In that case, one can question why A’ arguments are strong and

what are the reasons and measurements an AI system adopted to judge that A' arguments are strong, and whether an AI system had considered the ethical and moral principles in the decision that provided.

In conclusion, it can be asserted that understanding the fundamental causes of preferences can help in enhancing the use of AI systems, including AI-GC systems, in arbitration. Therefore, in the meantime, the arbitrators are advised to use these systems as a medium or tool to get inspiration and insight on writing the arbitral award in a formal, creative and genuine manner, but they should still be the ones doing the actual writing, reasoning and analysis, because of two main reasons. First, human capacity and intelligence regarding reasoning and common sense are far superior to AI systems, including AI-GC systems [34] [35]. Second, a lack of appropriate reasoning can result in a refusal to enforce or set aside an arbitral award in Malaysia for violating public policy.

6 Conclusion

Developments in AI technologies and the prompt transformation brought by such technologies lead to several critical questions and dilemmas. This article examined the legal and ethical conundrums of using AI-GC systems in arbitration. It found that although AI systems, including AI-GC systems, would play a vital role in arbitration and the legal industry, there are several legal and ethical conundrums in using those systems in arbitration. Therefore, several recommendations and insights have been provided throughout this article to enlighten not only the users, such as arbitrators but also the Malaysian lawmakers. Considering them will bring added value to the reputation of Malaysia as an advanced jurisdiction and significantly enhance the greater good of using AI technologies in arbitration industry.

References

1. Hussain, M. A., Labanieh, M. F., Mahdzir, N., Sulaiman, N. & Bawazir, O. S. A. The Potential Prospect Of Artificial Intelligence (AI) In Arbitration From The International, National, And Islamic Perspectives. *Journal of International Studies*. 2023. 19(1), 95-122.
2. Schuett, J. A legal definition of AI. 2019. Retrieved from https://www.researchgate.net/profile/Jonas-Schuett/publication/336198524_A_Legal_Definition_of_AI/links/5e20599a458515ba208b9e4c/A-Legal-Definition-of-AI.pdf
3. McCarthy, J. What is artificial intelligence. 2007. Retrieved from <https://www.diochnos.com/about/McCarthyWhatisAI.pdf>
4. Calo, R. Artificial intelligence policy: a primer and roadmap. *UCDL Rev.* 2017. 51(1), 399-435
5. Rouse, M. ChatGPT. *Techopedia*. 2023. Retrieved from <https://www.techopedia.com/definition/34933/chatgpt>.
6. Hencz, A. Agents Of Change: Artificial Intelligence –AI Art and How Machines Have Expanded Human Creativity. 2023. *Artland Magazine*. Retrieved from <https://magazine.artland.com/ai-art/>.

7. Rose, J. A Judge Just Used ChatGPT to Make a Court Decision. *Vice*. 2023. Retrieved from <https://www.vice.com/en/article/k7bdmv/judge-used-chatgpt-to-make-court-decision>
8. Jamal, S. Pakistani judge uses ChatGPT to make court decision. 2023. *Gulf News*. Retrieved from <https://gulfnnews.com/amp/world/asia/pakistan/pakistani-judge-uses-chatgpt-to-make-court-decision-1.95104528>
9. Smith, A, Moloney, A., Schapiro, A.A. (2023). Are AI chatbots in courts putting justice at risk?. 2023. *NST*. Retrieved from <https://www.nst.com.my/opinion/columnists/2023/05/906525/are-ai-chatbots-courts-putting-justice-risk>.
10. Labanieh, M. F. & Hussain, M. A. The Innovative Use of Artificial Intelligence in the Arbitral Tribunal: Malaysian Perspective. *International Journal of Advanced Science and Technology*. 2020. 29(6s), 1681-1687.
11. Hill, C. DISCO unveils large language model chatbot. *Legal Technology*. 2023. Retrieved from <https://legaltechnology.com/2023/03/23/disco-unveils-large-language-model-chatbot/>
12. Jones, J. How to use ChatGPT to summarize a book, article, or research paper. *Zdnet*. 2023. Retrieved from <https://www.zdnet.com/article/how-to-use-chatgpt-to-summarize-a-book-article-or-research-paper/>
13. Kira Systems. *Kira Systems*. 2023. Retrieved from <https://kirasystems.com/how-kira-works/>
14. Danziger, S., Levav, J., & Avnaim-Pesso, L. Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences*. 2011. 108(17), 6889-6892.
15. Schmitz, A. J. Expanding access to remedies through e-court initiatives. *Buffalo Law Review*. 2019. 67(1), 89-105
16. eBRAM. *Ai Machine Translation Service*. eBRAM. 2023. Retrieved from https://www.ebram.org/ai_machine_tran_services.html
17. Murphy, E.A. Standards of Arbitrator Impartiality: How Impartial Must They Be-Lifecare International, Inc. v. CD Medical, Inc. *Journal of Dispute Resolution*. 1996. 7(2): 463-469.
18. Rashid, S.K. Some Contentious issues in International Commercial Arbitration. *International Islamic University Malaysia Law Journal*. 2005. 13, 147-166.
19. Alam, S. Commercial Arbitration: Factors Making It Lose out against Other Processes. *Middle East Journal of Business*. 2014. 9(2), 51-59.
20. Brekoulakis, S, Systemic Bias and the Institution of International Arbitration: A New Approach to Arbitral Decision-Making. *Journal of International Dispute Settlement*. 2013. 4(3), 553-585.
21. Nousia, K. P. Bias of arbitrators revisited. *Journal of Business Law*. 2018. 9(4), 1-29
22. Sun, C.L. "Arbitrators' Conflicts of Interest: Bias by Any Name. *Singapore Academy of Law Journal*. 2007. 19(1), 254-265.
23. Murphy, H. T. Artificial Intelligence Is Dreaming Up a Very White World. *Slate*. 2023. Retrieved from <https://slate.com/technology/2023/02/dalle2-stable-diffusion-ai-art-race-bias.html>
24. Obermeyer, Z., et al. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*. 2019. 366(6464), 447-453.
25. Osborne Clarke. *AI and the future of Dispute Resolution: Computer Says No?*. Osborne Clarke. 2021. Retrieved from <https://www.osborneclarke.com/insights/ai-future-dispute-resolution-computer-says-no#:~:text=But%20%22baked%20in%22%20biases%20in,those%20biases%20in%20its%20predictions>
26. Davies, S.C., et al. A computer program used for bail and sentencing decisions was labeled biased against blacks. It's actually not that clear. *Washington Post*. 2016. Retrieved from <https://www.washingtonpost.com/news/monkey-cage/wp/2016/10/17/can-an-algorithm-be-racist-our-analysis-is-more-cautious-than-propublicas/>

27. Thorn, P.D., Bostrom, N. Superintelligence: Paths, Dangers, Strategies. *Minds & Machines*. 2015. 25, 285–289.
28. Mishchenko, T. Meta shuts down public test of Galactica, its ‘AI for Science’ because it produced pseudoscientific papers. *Mezha Media*. 2021. Retrieved from <https://mezha.media/en/2022/11/21/meta-shuts-down-public-test-of-galactica-its-ai-for-science-because-it-produced-pseudoscientific-papers/>
29. Umawing, J. Now it's BlenderBot's turn to make shocking, inappropriate, and untrue remarks. *Malware Bytes*. 2022. Retrieved from <https://www.malwarebytes.com/blog/news/2022/08/its-now-blenderbots-turn-to-make-shocking-inappropriate-and-untrue-remarks>.
30. Moran, C. ChatGPT is making up fake Guardian articles. Here’s how we’re responding. *The Guardian*. 2023. Retrieved from <https://www.theguardian.com/commentis-free/2023/apr/06/ai-chatgpt-guardian-technology-risks-fake-article>
31. Delouya, S. I asked ChatGPT to do my work and write an Insider article for me. It quickly generated an alarmingly convincing article filled with misinformation. *Business Insider*. 2022. Retrieved from <https://www.businessinsider.com/i-asked-chatgpt-to-write-insider-story-it-was-convincing-2022-12>
32. Greenberg, S, Kee, C. & Weeramantry, J.R. *International Commercial Arbitration: An Asia Pacific Perspective*. 2012. United States of America, New York: Cambridge University Press.
33. Burrell, J. How the machine ‘thinks’: Understanding opacity in machine learning algorithms. *Big Data & Society*. 2016. 3(1), 1-12.
34. Knight, W. Tougher Turing Test Exposes Chatbots’ Stupidity. *Technology Review*. 2016. Retrieved from <https://www.technologyreview.com/2016/07/14/7797/tougher-turing-test-exposes-chatbots-stupidity/>
35. Hao, K. AI still doesn’t have the common sense to understand human language. *Technology Review*. 2020. Retrieved from <https://www.technologyreview.com/2020/01/31/304844/ai-common-sense-reads-human-language-ai2/>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

