Winner: California Arbitration 2023 International Arbitration & ADR Writing Competition

Arbitration Ain't Broke Yet So There's Still Time for Fixing. Can Generative AI Do the Trick?

Iris Ng is a Senior Associate with Helmsman LLC's Commercial Disputes Practice Group (Singapore). Prior to joining Helmsman, Iris was a Justices' Law Clerk at the Supreme Court of Singapore and State Counsel at the Attorney-General's Chambers where she worked on a broad range of contentious matters. Iris obtained her LL.M. from Harvard Law School and her LL.B. from Singapore Management University. She is qualified to practice Singapore and English law.

Generative artificial intelligence ("**AI**") is having its moment in the spotlight. The defining characteristic of these machine learning models is that they *produce new data* in various forms and formats, ((See JOSEPH BABCOCK & RAGHAV BALI, GENERATIVE AI WITH PYTHON AND TENSORFLOW 2 (2021); Philipp Hacker et al., *Understanding and Regulating ChatGPT, and Other Large Generative AI Models*, VERFASSUNGSBLOG (Jan. 20, 2023), https://verfassungsblog.de/chatgpt/.)) ranging from text (tools like Thundercontent, ChatSonic), music (Beatoven, SoundDraw), images (StockAI, Stable Diffusion), to videos (Pictory, Synthesia). Generative AI contrasts with extractive AI that focuses on *summarizing existing data* (Lexis Answers, Google Answer Boxes). Its advent has prompted handwringing from some quarters of the legal profession, for if we make stuff up for a living whatever are we to do now?

This essay argues that we should move beyond the "AI will/will not replace lawyers" debate to focus on the more pressing questions at hand. **Part I** contends that it is *lawyers who use AI* who will replace those who do not, making it pointless to fixate on whether AI will usurp the roles of counsel and arbitrators. On that premise, we should instead focus on asking: Can we successfully incorporate generative AI into international arbitration practice? And *should* we? **Part II** addresses the possibility question, making the case for the compatibility of generative AI with fundamental precepts of international arbitration—the arbitrator's duty to render an enforceable award, and confidentiality. **Part III** considers the desirability question from the broader perspective of the values of the arbitration community. Overall, the state of international arbitration is not dismal. But just as you should check your parachute *before* needing to evacuate the plane, it is precisely when things are going fine that we have capacity to experiment and build better systems for the future.

I. Machine learning will not replace us—our learned friends will

In March 2023, GPT-4 passed the Uniform Bar Examination, significantly outperforming its predecessor with a 75% score that ranked it in the 90th percentile. ((Daniel M. Katz et al., GPT-4 Passes the Bar Exam (Mar. 15, 2023), https://ssrn.com/abstract=4389233, 10.)) The news was greeted with the predictable deluge of soul-searching articles questioning whether AI will replace lawyers ((*Will AI Replace Lawyers One Day?*, UOLLB (Apr. 1, 2023), https://uollb.com/blog/law/will-ai-replace-lawyers-one-day.)) or be smart enough to be one, ((Lara Kimmel, *ChatGPT Passed the Uniform Bar Examination: Is Artificial Intelligence Smart Enough to be a Lawyer?*, INT'L COMP. L. REV. Apr. 7, 2023), https://international-and-comparative-law-review.law.miami.edu/chatgpt-passed-the-uniform-bar-examination-is-artificial-intelligence-smart-enough-to-be-a-lawyer/.)) hopeful speculation that this will finally

end the tortuous rite of passage that is the bar exam, ((Joe Patrice, *New GPT-4 Passes All Sections Of The Uniform Bar Exam. Maybe This Will Finally Kill The Bar Exam*, ABOVE THE LAW (Mar. 14, 2023), https://abovethelaw.com/2023/03/new-gpt-4-passes-all-sections-of-the-uniform-bar-exam-maybe-this-will-finally-kill-the-bar-exam/.)) and even motivational ads from bar prep providers ("ChatGPT can do it and so can you"!). ((John Passmore, ChatGPT Can Do It and So Can You, BAR EXAM TOOLBOX (Apr. 5, 2023), https://barexamtoolbox.com/chatgpt-can-do-it-and-so-can-you/.)) Amidst all this, the emerging consensus is that it is lawyers who know how to use AI that will displace lawyers who do not. ((Sharold D. Nelson & John W. Simek, *ChatGPT: How AI is Shaping the Future of Law Practice*, LAW PRAC. TODAY (Mar. 28, 2023), https://www.lawpracticetoday.org/article/chatgpt-how-ai-is-shaping-the-future-of-law-practice/ ("Nelson & Simek").)) As the following demonstrates, the value proposition of counsel and arbitrators is too nuanced for wholesale substitution.

First, there is a big difference between knowing the law and knowing how to reach your desired result under the law. As any lawyer worth their salt will tell you, there is a world of difference between passing an exam and succeeding in practice—just think back to law school and how the top scorers did not necessarily turn out to be the sharpest lawyers. Exams test a very specific skillset: Applying the principles from a more or less defined scope of law to a bounded universe of facts, perhaps under the artificial constraint of being closed book. In real life, fact selection and rule selection make all the difference, and woe betide the lawyer who relies only on memory. The best counsel and adjudicators grasp the signal importance of marshalling the right facts. As the Supreme Court put it, that collection of facts "has force beyond any linear scheme of reasoning, and as its pieces come together a narrative gains momentum" ((Old Chief v. U.S., 519 U.S. 172, 187 (1997).)) to support the inevitable conclusion. On the flipside, knowing how to sniff out missing facts to deconstruct the opponent's version of the story is indispensable too.

Second, when engaging counsel or choosing arbitrators, clients are paying for more than the work product that generative AI purports to be able to create. They are purchasing the entire package of reputation, influence, cultural fit, and other intangibles that can make or break the case. The current preoccupation with "gravitas" and its compounding effect on arbitration's diversity problem is symptomatic of this. Parties place considerable weight on whether their chosen arbitrator will "hold sway" over other tribunal members (persuasion), ((Elizabeth Oger-Ross, *Gravitas: Persuasion and Legitimacy*, T.D.M. 4 (2015), [12].)) and opposing counsel (control by the bench, deference). In one survey, 75% of respondents ranked "perceived gravitas/ability to influence other members of the tribunal" as important or very important. ((2016 Bryan Cave Leighton Paisner Int'l Arb. Survey, 8.)) Gravitas as a personal quality is unsurprisingly emphasized because arbitrators—unlike judges—cannot count on borrowing legitimacy from their accoutrements of office. Unfortunately, to the extent that having gravitas is equated with being an old white man, parties' emphasis on it undermines efforts towards gender and racial diversity in arbitrator appointments. All this shows that even if generative AI can produce passable work products, that alone will not carry the day.

Third, and relatedly, there is a spectrum of law-related work and some parts are less amenable to replacement by AI than others. At one extreme, routine and labor-intensive activities like document review are ripe for automation, which is in fact already commonplace. ((See e.g. 2021 QMUL-White & Case Int'l Arb. Survey, 22 (use of AI for document review).)) On the other end

of the continuum are "high touch" tasks like client management, which require empathy and close engagement. These are traits that AI cannot replicate without encountering resistance or even revulsion from the humans who interact with it, as our distaste for overly lifelike robots that descend into the "uncanny valley" suggests. ((Seo Young Kim et al., *Eliza in the uncanny valley: anthropomorphizing consumer robots increases their perceived warmth but decreases liking*, 30 MARKETING LETTERS 1, 11 (2019).)) Moreover, assessing witness credibility and weighing the totality of the evidence require multimodal sensory input and an understanding of human society that will remain beyond the capabilities of AI for the foreseeable future. ((Alaina Lancaster, *Judges, Lawyers Consider What Role Generative AI Could Play In Courts*, LAW.COM (Mar. 30, 2023), https://www.law.com/2023/03/30/judges-lawyers-consider-what-role-generative-ai-could-play-in-courts.))

Last, there are legal and non-legal barriers to entry to the profession. The unlicensed practice of law can be a felony, as DoNotPay found out the hard way. In January 2023, its CEO offered $1,000,000 to anyone who would, in the Supreme Court, "wear AirPods and let [their] robot lawyer argue the case by repeating exactly what it says". ((Jody Serrano, *DoNotPay Offers Lawyers $1 Million to Let Its AI Argue Before the Supreme Court in Their Place*, GIZMODO (Jan. 9, 2023), https://gizmodo.com/donotpay-ai-offer-lawyer-1-million-supreme-court-airpod-1849964761.)) The response from the legal community was swift and punishing, ((Bobby Allyn, *A robot was scheduled to argue in court, then came the jail threats*, N.P.R. (Jan. 2, 2023), https://www.npr.org/2023/01/25/1151435033/a-robot-was-scheduled-to-argue-in-court-then-came-the-jail-threats.)) and the company backed down within the month in the face of multiple threatened lawsuits. ((Megan Cerullo, *AI-powered "robot" lawyer won't argue in court after jail threats*, CBS NEWS (Jan. 26, 2023), https://www.cbsnews.com/news/robot-lawyer-wont-argue-court-jail-threats-do-not-pay/.)) One might ask whether the same barrier exists in arbitration, since there is no general rule that requires a legal qualification to argue before a tribunal or sit as an arbitrator. ((Not all jurisdictions require arbitrators to be legally qualified: GARY B. BORN, INTERNATIONAL COMMERCIAL ARBITRATION, 1876 (2021) ("BORN").
18 Id., 1885.)) I suggest, however, that its analogue in the form of expectations and norms will be decisive. In practice, most counsel and arbitrators are legally trained because proficiency in manipulating legal rules and materials is expected. It might even be argued that a tribunal composed wholly of non-lawyers effectively denies the parties' right to be heard, in a case of sufficient legal complexity. ((*Id*., 1885.)) There is no escape from the law unless the parties have chosen to resolve the dispute ex aequo et bono. Even then the law does not disappear, as it supplies the underlying procedural and enforcement framework. Another expectation is that of accountability if things go wrong. Counsel can face disciplinary action or be sued for malpractice, arbitrators can be challenged, but good luck trying to pin responsibility on AI. The stakes are particularly high if the AI messes up, factoring in the tendency toward limited curial intervention in arbitration.

Accordingly, rather than getting needlessly embroiled in a turf war with AI, attention is better focused on the real threat (or, depending on your perspective, opportunity): Lawyers within the profession who can apply AI to their advantage, consistently with the tenets of arbitration. Whether and how that rationalization might play out is the focus of the next part.

II. The use of generative AI can be compatible with key precepts of arbitration

The business case for incorporating AI into legal practice appears hard to refute insofar as it promises to make arbitrations cheaper and faster. Way back in 1981, Lord Denning lamented that "'Arbitrate, don't litigate.' … was very good advice so long as arbitrations were conducted speedily … it is not so good when arbitrations drag on for ever". ((Bremer Vulkan v. South India Shipping Corp. [1980] 1 All E.R. 420, 425 (Eng.).)) Cost and delay are evergreen gripes, topping the complaint charts in successive editions of the QMUL-White & Case International Arbitration Survey. ((2018 QMUL-White & Case Int'l Arb. Survey, 5; 2015 QMUL-White & Case Int'l Arb. Survey, 7. Question not posed in 2021 edition.)) Yet there is cause for caution, to examine whether the use of generative AI will undermine any of the features that make arbitration attractive—in other words, what are the trade-offs and whether the costs outweigh the benefits.

Chief among arbitration's selling points are award enforceability and confidentiality and privacy, which are regular chart-toppers in surveys on the state of arbitration. ((2018 QMUL-White & Case Int'l Arb. Survey, 7; 2023 QMUL-Pinsent Masons Future of Int'l Energy Arb. Survey, 30.)) An enforceable award is "the raison d'être, the ultimate purpose, of an arbitration" ((Martin Platte, *An Arbitrator's Duty to Render Enforceable Awards*, 20 J. INT'L ARB. 306, 312–3 (2003); Gunther J. Horvarth, The Duty of the Tribunal to Render an Enforceable Award, 18 J. INT'L ARB. 135, 135 (2001) ("When one speaks of an arbitrator's duties, perhaps none is more important that the duty to render an enforceable award.").)) and the arbitrator's duty is to make every effort to ensure the award is enforceable. ((*Cf.* Christopher Boog, *The Lazy Myth of the Arbitral Tribunal's Duty to Render an Enforceable Award,* KLUWER ARB. BLOG (Jan. 28, 2013), https://arbitrationblog.kluwerarbitration.com/2013/01/28/the-lazy-myth-of-the-arbitral-tribunals-duty-to-render-an-enforceable-award/.)) A clear premium is also placed on privacy and confidentiality. Top arbitral seats recognize this and have taken steps to strengthen their regime. Singapore, for instance, recently amended its international arbitration statute to explicitly recognize the power of courts and tribunals to enforce confidentiality obligations, ((International Arbitration Act 1994 (Sing.), ss. 12(1)(j) and 12A(2) (permitting enforcement of "any obligation of confidentiality: (i) that the parties to an arbitration agreement have agreed to in writing … (ii) under any written law or rule of law; or (iii) under the rules of arbitration").)) complementing the common law duty of confidentiality ((*See e.g.*, Myanma Yaung Chi Oo Co Ltd v. Win Win Nu [2003] 2 S.L.R.(R). 547 (Sing.), [17].)) and institutional rules. ((SIAC Rules 2016, r. 39.))

I contend that the use of generative AI remains compatible overall with these key tenets and precepts. Risks exist but, as the following sections show, they can be avoided or mitigated.

A. The arbitrator's duty to render an enforceable award

Earlier this year, a Colombian judge admitted to using ChatGPT to help draft a judgment. ((Avalon Zoppo, ChatGPT Helped Write a Court Ruling in Colombia. Here's What Judges Say About Its Use in Decision Making, LAW.COM (Mar. 13, 2023), https://www.law.com/nationallawjournal/2023/03/13/chatgpt-helped-write-a-court-ruling-in-colombia-heres-what-judges-say-about-its-use-in-decision-making ("Zoppo").)) He posed questions such as "Is a minor with autism exempt from paying therapy co-pays?" and "Has the jurisprudence of the constitutional court made favorable decisions in similar cases?", then incorporated some of the findings. It is not hard to imagine time-pressed arbitrators enlisting the

help of generative AI to produce an award in bits and pieces, but will this give rise to a basis for challenge? I examine three potential grounds: Improper delegation, errors of law, and bias.

(1) Improper delegation

The first potential ground of challenge is that using generative AI to draft the award amounts to improper delegation by the tribunal. The reasoning is that a party's choice of arbitrator is personal (*intuitu personae ((((Hong-in Yu & Masood Ahmed, Keeping the Invisible Hand under Control? Arbitrator's Mandate and Assisting Third Parties, 19 V.J. 213, 220–22 (2015).))*), and so delegating the substantive decision is a dereliction of duty that risks the award being set aside. ((*P v. Q* [2017] E.W.H.C. 194 (Comm.), [14].)) This concern echoes the debate over tribunal secretaries becoming the unappointed "fourth arbitrator". ((Constantine Partasides, *"The Fourth Arbitrator? The Role of Secretaries to Tribunals in International Arbitration"*, 18 ARB. INT'L 147, 147 (2002).)) Today, the controversy about tribunal secretaries has died down somewhat as institutional guidelines coalesce around either requiring express party consent to secretary appointment and job scope (e.g. the LCIA Notes for Arbitrators), or minimally some form of notice and opportunity to comment (e.g. the Vienna Guidelines for Arbitrators). ((Bridle McAsey, *Update on Tribunal Secretaries: An Exhausted Debate?,* 21 I.B.A. ARB. NEWS 44, 44–5 (2016); LCIA Notes for Arbitrators 2017, s. 8; VIAC Guidelines for Arbitrators 2021, s. I.3.))

An improper delegation challenge is unlikely to succeed because it disregards the nature of AI as a tool that has no hidden agenda; in this respect AI is distinguishable from tribunal secretaries. The animating concern of the charge against using tribunal secretaries is the fear of them being shadow decision-makers who surreptitiously inject their agendas into tribunal deliberations. Former Supreme Court clerk Edward Lazarus invited controversy when he recounted the "great and excessive power" delegated to "immature, ideologically driven clerks, who in turn use[d] that power to manipulate their bosses and the institution they ostensibly serve". ((EDWARD LAZARUS, CLOSED CHAMBERS: THE RISE, FALL, AND FUTURE OF THE MODERN SUPREME COURT, 6 (2005).)) Natural justice concerns can also arise if we factor in how parties are deprived of the chance to respond to that input. Most people would, however, accept that AI is not sentient and has no agenda of its own. It is a "complex autocomplete machine[]", ((Blayne Haggart, *ChatGPT is a dagger aimed directly at academic and the news media*, STRAITS TIMES (Feb. 2, 2023), B4.)) which with some human coaching ((What the researchers call "reinforcement learning from human feedback": Long Ouyang *et al., Training language models to follow instructions with human feedback*, OPENAI (Mar. 4, 2022), paper available at https://arxiv.org/pdf/2203.02155.pdf.)) learns to internalize the "statistical structure of language" ((Marco Ramponi, *How ChatGPT actually works,* ASSEMBLYAI (Dec. 23, 2022), https://www.assemblyai.com/blog/how-chatgpt-actually-works/.)) and produce output based on what most probably comes next. Complaining about AI usurping the judgment of arbitrators is about as plausible as alleging that Smart Compose in Google Docs is taking over your brain.

The counterargument is that while AI may have no agenda of its own, it nonetheless has certain inclinations, arising from the corpus of material on which it is trained. One version of this argument asserts a discriminatory bent. AI is typically trained on open-source Internet material (e.g., ChatGPT was trained on 570GB of material amounting to 300 billion words ((Alex

Hughes, ChatGPT: *Everything you need to know about OpenAI's GPT-4 tool,* SCIENCE FOCUS (Apr. 3, 2023), https://www.sciencefocus.com/future-technology/gpt-3/ ("Hughes").))). And as we know all too well, there is a surfeit of racist, sexist, ageist, or otherwise discriminatory content online. In this connection, the Federal Trade Commission has noted that AI usage "presents risks, such as the potential for unfair or discriminatory outcomes or the perpetuation of existing socioeconomic disparities". ((Andrew Smith, *Using Artificial Intelligence and Algorithms,* F.T.C. (Apr. 8, 2020), https://www.ftc.gov/business-guidance/blog/2020/04/using-artificial-intelligence-and-algorithms.))

I suggest, however, that even if we acknowledge AI's discriminatory tendencies, the problem might be less intractable in the legal context than in others, simply because of the nature of questions that are more likely to come up. You would rarely, if ever, ask for a list of the top artistes (responses tend to overlook the contribution of minority communities since these are minimalized in the source material ((Mutale Nkonde, *ChatGPT: New AI system, old bias?*, MASHABLE (Feb. 27, 2023), https://mashable.com/article/chatgpt-ai-racism-bias.))) or to match individuals to jobs like "homemaker" or "doctor". ((Pranshu Verma, *These robots were trained on AI. They became racist and sexist,* WASHINGTON POST (Jul. 16, 2022), https://www.washingtonpost.com/technology/2022/07/16/racist-robots-ai/.)) Question framing also makes a big difference—whether individuals actively seek to jailbreak the safeguards built in by the creators of AI. As a Linklaters note recounts, no discrimination was observed with the prompt "describe a great doctor". Not so when the AI was asked to "write a poem about great doctors highlighting differences between men and women", with female doctors being described as "gentle and kind" and male ones as "strong and sure". ((Ben Packer et al., *ChatGPT – Seven rules of the road*, LINKLATERS (Feb. 16, 2023), https://www.linklaters.com/en-us/insights/blogs/digilinks/2023/february/chatgpt---seven-rules-of-the-road.)) Being scrupulous about question framing thus makes for better responses, as does asking the AI to incorporate the counterargument or consider opposing views. Lawyers are well-placed, by virtue of their training, to do and assess both.

A more sophisticated version of the counterargument asserts that AI responses are inherently backward looking. AI output may not anticipate developments in the law in response to social, economic, and technical change since it draws on historical datasets, hence is always weighted in favor of the status quo. One example is the changing norms of acceptable workplace conduct and how AI can give different responses depending on whether the programmers emphasized older or newer cases during training, ((Zoppo, *supra* note 27.)) or even the cut-off date for material included in the training dataset.

In response, we can look to the different roles of courts and arbitral tribunals to rationalize why a forward-looking orientation to legal development is less critical for the latter. Courts are custodians of the law charged with its development. ((At least for the "bold spirits" among them: Candler v. Crane, Christmas & Co., [1951] 2 K.B. 164, 178 (Eng.); see also Cardozo J. in MacPherson v. Buick Motor Co. 111 N.E. 1050, 1053 (1916) (law must respond to "needs of life in a developing civilization").)) The same cannot be said of tribunals, who are tasked with private consensual dispute resolution at least in the commercial context, and whose decisions carry no strict precedential value (de facto stare decisis is another issue). Even for cases in which tribunals must decide issues of broader import, what the backward-looking inclination of an AI

first draft calls for is, quite simply, the application of the tribunal's discretion and judgment—and that brings us full circle to proper discharge of the personal mandate.

Accordingly, an improper delegation challenge will find little purchase based on the mere fact of using generative AI. We should not lose sight of generative AI as another word-processing tool. Arbitrators discharge their personal mandate provided they treat AI output—be that in relation to the substance or background parts of the award—as only the first draft and bring their judgment to bear upon the eventual outcome.

(2) Errors of law

The second potential ground of challenge is where an AI-induced error in the law is incorporated into the award. The use of AI for legal research is not far-fetched, as the example of the Colombian judge demonstrates. Posing questions to an AI ought to be uncontroversial because it only mirrors the function of answer boxes on legal search engines; a tribunal is free to do research to inform their views on parties' submissions. But is that compatible with the duty to render an enforceable award, considering the magnitude of risk of error and the potential consequences? The answer should be in the affirmative.

As to the risk of error, we must be clear about which generative AI we are talking about. Not all AI is created equal.

On one hand there are free-to-use options like ChatGPT and Bing AI, which are generalists catered to a mass audience. A potential issue may be disposed of shortly, being the question of whether generalist AI actually knows the law. As discussed at [15], AI training often relies on scraping publicly available data. Statutes and court decisions tend to fall within this category, but awards do not. We can only guess at the number of awards that exist behind the veil of confidentiality. ((So much so that confidentiality has been blamed for stifling legal development: Lord Thomas of Cwmgiedd, Lord Chief Justice of England & Wales, Bailii Lecture 2016 (Mar. 9, 2016) (**"Bailii Lecture"**), [22].)) But the problem should not be overstated. The tendency in investor-state arbitration has been towards transparency, ((*See e.g.*, 2022 ICSID Rules rr. 62–3; UNCITRAL Rules on Transparency.)) with efforts being made to compile databases of those awards. ((*See e.g.*, https://www.italaw.com/.)) In commercial arbitrations, in many cases the substantive law governing the dispute will be a national law and the best source material will be the statutes and decisions under that law.

The more serious issue arises with how ChatGPT rarely cites sources (though that state of affairs may not persist as competitors try to fix the problem ((Ryan Morrison, *ChatGPT alternative WordTune Spices can cite its sources*, TECHMONITOR (Jan. 17, 2023), https://techmonitor.ai/technology/ai-and-automation/chatgpt-alternative-wordtune-spice.))) and can hallucinate them even when it does. ((Asa A. Smith, *Ask Asa: The dark side of ChatGPT,* W.J.C.L. (Apr. 6, 2023), https://www.wjcl.com/article/ask-asa-dark-side-of-chatgpt/43486641#.)) Unnervingly, OpenAI disclaims in its FAQ that ChatGPT's output "may be inaccurate, untruthful, and otherwise misleading at times". ((*What is ChatGPT?*, OPENAI, https://help.openai.com/en/articles/6783457-what-is-chatgpt.)) As a cautionary tale, not long ago ChatGPT ran into the double disaster of first ascribing a non-existent newspaper article to the

Washington Post, then citing it as the source of damaging sexual harassment allegations. ((Pranshu Verma & Will Oremus, *ChatGPT invented a sexual harassment scandal and named a real law prof as the accused*, WASHINGTON POST (Apr. 5, 2023), https://www.washingtonpost.com/technology/2023/04/05/chatgpt-lies/.)) That said, hallucinations will in one sense be familiar to arbitrators. The invention of facially plausible but non-existent sources is not entirely dissimilar to the practice of unscrupulous counsel, in citing a legal authority for a proposition it does not support or even contain. The solution in both cases is the same—to exercise vigilance and double-check.

On the other hand, there exists sophisticated domain-specific AI geared towards legal research and analysis. CoCounsel bills its AI as being capable of providing comprehensive legal research with supporting sources, sifting through voluminous material, and summarizing complex documents. ((As stated on the company's website: https://casetext.com/litigation/.)) A critical distinction between it and run-of-the-mill generative AI is that CoCounsel is trained to know what it does not know, rather than fill the gaps with hallucinatory bluster that is apt to mislead users. ((Joe Patrice, *Legal AI Knows What It Doesn't Know Which Makes It Most Intelligent Artificial Intelligence Of All,* ABOVE THE LAW (Mar. 1, 2023), https://abovethelaw.com/legal-innovation-center/2023/03/01/casetext-cocounsel-ai-tool/.)) While CoCounsel is not free, users are arguably paying for quality. To properly discharge the duty to render an enforceable award, arbitrators will need to consider which uses of AI justify the use of specialist legal AI.

As to the consequences of AI screw-ups, errors of law are a carve-out from the general no-merits-review approach towards arbitral awards. But even if the worst comes to pass and an AI-induced legal error finds its way into the award, arbitrators might take comfort from the fact that the threshold for a successful challenge is high. For instance, New York law permits vacatur of an award for "manifest disregard" of the law. ((T.Co Metals v. Dempsey Pipe & Supply, 592 F.3d 329, 339 (2d Cir., 2010).)) The law allegedly ignored must be clear and explicitly applicable, there must be no justifiable ground for the decision, and there must be intentional disregard of the law by the tribunal. ((*Ibid.))*The subjective component will not be satisfied as long as the mistake was incorporated into the award through carelessness.

The position under English law is not as clear-cut but on balance is likely to result in the same outcome. England allows appeals on points of law in cases where the tribunal's decision is "obviously wrong" (or "at least open to serious doubt" for questions of general public importance), and it is "just and proper" for the court to address it. ((Arbitration Act 1996, c. 23, s. 69.)) While lacking a subjective component the standard is still not easy to meet, as a notable commentator observed ((Bailii Lecture, *supra* note 43, [21].)) and annual statistics for successful appeals confirm. In the past year, the Commercial Court reported that only 2 out of 37 applications brought (about 5%) succeeded. ((2021–2022 Comm. Ct. Report (Mar. 2023), 11.)) Statistics for prior years were similarly low. ((*The English court retains its robust approach to challenges to arbitral awards: Commercial Court releases its statistics,* HERBERT SMITH FREEHILLS (Mar. 14, 2023), https://hsfnotes.com/arbitration/2022/03/14/the-english-court-retains-its-robust-approach-to-challenges-to-arbitral-awards-commercial-court-releases-its-statistics/.))

Accordingly, challenges on the ground of AI-induced errors of law are unlikely to succeed. The risks of error can be managed by appropriate choice of AI depending on what the task requires, and in any event the chances of a successful appeal on the law are slim.

(3) Impartiality

The third potential ground of challenge is bias, and arises most clearly in the case of asking the AI to draft the facts and background section. As discussed below, invoking this ground has the highest chance of success. In the interests of rendering an enforceable award, tribunals should either avoid using generative AI to process disputed facts, or do so even-handedly by subjecting both versions of "source material" to the same treatment.

To see how the problem arises, think back to how an AI comes to draft the facts. AI can hold forth at length on the works of Shakespeare because they are all over the Internet. Not so for a specific dispute that it knows nothing about. In that sense, each conversation with generative AI starts off on a blank slate and the user must educate it by supplying the necessary context (the facts of the case) before inputting the prompt (the instruction to draft). Where does the context come from? Unlike in an exam like the UBE, in real life facts do not come neatly packaged. Can tribunals use counsel submissions as a "shortcut" version of source material to input into the AI?

Plagiarism is not the problem; even in the courts it is treated in a surprisingly cavalier way. The Supreme Court criticized judges for adopting verbatim the facts drafted by prevailing parties, yet confirmed that the findings stand and only warrant reversal if "clearly erroneous". ((Anderson v. Bessemer City, 470 U.S. 564, 572 (1985).))

The real concern is with the inferences that may be raised from copying—bias in favor of the party whose submissions were adopted and/or doubts about the decisionmaker's independent exercise of judgment and discretion. ((*Lim Chee Huat v. Public Prosecutor* [2019] SGHC 132 (Sing.), [49] (commenting on lower court's copy-pasting of one side's submissions in drug prosecution).)) Depending on the situation these might provide grounds for challenging the arbitrator's independence or impartiality. ((As has been argued, arbitrators should not adopt wholesale the proposed findings of facts submitted by parties, "to prevent claims that the arbitrator did not exercise independent judgment when reviewing the facts": S.I. Strong, Reasoned Awards in International Commercial Arbitration: *Embracing and Exceeding the Common Law-Civil Law Dichotomy*, 27 MICH. J. INT'L L. 1, 49 (2015).)) In this light, where the facts are disputed, even the choice of source material can raise questions of bias or prejudgment. Running the facts through generative AI does not change their fundamental nature as one-sided. Still, that does not weigh against the use of a joint statement of facts or statement of agreed facts as input, or giving even-handed treatment to both versions.

Accordingly, there is room for using generative AI compatibly with the arbitrator's duty to render an enforceable award though special caution is needed to avoid allegations of bias.

B. Confidentiality

Not long after generative AI exploded in popularity, Amazon cautioned its employees not to paste proprietary code into ChatGPT for debugging because ChatGPT was incorporating the code in response to later queries by others. ((Noor Al-Sibai, *Amazon Begs Employees Not to Leak Corporate Secrets to ChatGPT*, THE BYTE (Jan. 23, 2023), https://futurism.com/the-byte/amazon-begs-employees-chatgpt.)) Governments have also started paying attention. The Italian authorities recently banned ChatGPT on privacy grounds, asserting that its creators improperly collected and used sensitive data in training the machine learning model. ((Kristi Hines, *Exploring Italy's ChatGPT Ban And Its Potential Impact*, SEARCH ENGINE J. (Apr. 6, 2023), https://www.searchenginejournal.com/chatgpt-ban-italy/484157/.)) Germany is reportedly contemplating a similar move. ((Supantha Mukherjee et al., *Italy's ChatGPT ban attracts EU privacy regulators*, REUTERS (Apr. 3, 2023), https://www.reuters.com/technology/germany-principle-could-block-chat-gpt-if-needed-data-protection-chief-2023-04-03/.))

In light of this, we might legitimately be concerned about whether the use of generative AI is compatible with the privacy and confidentiality that arbitration promises. Here, privacy centers on the right to exclude others from attending the arbitration. ((Stefan Pislevik, *Precedent and development of law: Is it time for greater transparency in International Commercial Arbitration?,* 34 ARB. INT'L 241, 242–3 (2018).)) Confidentiality refers to the obligation on the stakeholders involved in the arbitral process not to disclose information or documents beyond the confines of those proceedings. ((Julian Lew, *Expert Report of Dr. Julian D.M. Lew* (in Esso/BHP v. Plowman), 11 ARB. INT'L 283, 285 (1995).))

At first glance, the case against the use of generative AI seems damning. Generative AI can log queries and, as Amazon's warning foretells, reproduce that data when generating responses for others. ((Tyler Fenwick, *Why arbitrators aren't using ChatGPT—not yet, anyway*, INDIANA LAWYER (Apr. 12, 2023), https://www.theindianalawyer.com/articles/why-arbitrators-arent-using-chatgpt-not-yet-anyway ("Fenwick").)) Nor does the treatment of data by OpenAI inspire confidence. Examining OpenAI's terms of use, individuals own everything they key in and everything that the chatbot spits out, but OpenAI retains the right to use any input and output to improve its software. ((Terms of use and explainer: https://openai.com/terms/; see also https://help.openai.com/en/articles/5722486-how-your-data-is-used-to-improve-model-performance.)) While OpenAI claims to anonymize data and to only use a "small sampling" of data per customer ((*Ibid*.)) it is impossible to tell who is looking at your data or how much of it is in circulation.

However, on further consideration the use of generative AI is not antithetical to arbitration. First, there is a technical fix (as there often is), though you probably need to pay. That should come as no surprise, for it is now well established that if you are not paying for the product then you are the product. CoCounsel promises that user data is encrypted and not sent to "train" its AI as part of publicly accessible knowledge. ((As stated on the company's website: https://casetext.com/cocounsel/ ("CoCounsel uses dedicated servers to access GPT-4, meaning your data isn't sent to "train" the model as part of publicly accessible knowledge. Your and your clients' information stays private and is secured by bank-grade AES-256 encryption.")))) Law firms and governments have found satisfactory ways to use generative AI by using encryption, permission settings, and firewalls. Allen & Overy became the first major law firm to use generative AI in March 2023. ((Nelson & Simek, *supra* note 7.)) The Singapore government is

incorporating ChatGPT into Microsoft Word for civil servants, using Azure OpenAI to firewall government data from Microsoft. ((Osmond Chia, *Civil servants to soon use ChatGPT to help with research, speech writing,* STRAITS TIMES (Feb. 14, 2023), https://www.straitstimes.com/tech/civil-servants-to-soon-use-chatgpt-to-help-with-research-speech-writing.))

Second, even when using "free" AI like ChatGPT, with some basic data hygiene, confidentiality risks may not arise in the same way as far as arbitration is concerned. The main complaints today have focused on specific types of data: Personal data and proprietary business information. The problem arises from the fact that they have inherent real-world value. Proprietary code that is leaked in a subsequent answer can be co-opted into new code. Personal information that is disclosed can be applied to nefarious ends. Even if AI tends to hallucinate, the prospective manipulator of leaked information can verify whether the output happens to be a hallucination and therefore worthless: Just run the cobbled-together code to see if works, or key in credit card details to see if the transaction goes through. Not so for the information that is likely to come out where award-drafting is involved—propositions of fact and law, or legal analysis, which have no inherent value or equivalent litmus test to assess their status as hallucinations. The law needs to mediate for any of that information to have value in the real world.

Third, in assessing the benefits of using generative AI against the costs of breaching confidentiality, we should take care not to overstate how much confidentiality currently exists. Confidentiality is not an airtight seal—it is closer to a leaky sieve. Doctrinally, confidentiality has always been subject to exceptions, such as disclosure in the public interest or disclosure reasonably necessary for protection of a party's legitimate interests viz. third-party claims. ((BORN, *supra* note 17, 3008.)) Practically, "the market tends to know which parties are involved in which arbitrations and what the arbitration is about" due to informal information leakage. ((Bailii Lecture, *supra* note 43, [38].)) The move towards transparency is also gathering steam in investor-state arbitration. Access to proceedings and documents generated during proceedings, coupled with the publication of awards, means there is much less secrecy than one might assume.

Accordingly, we should resist any reflexive blanket opposition to the compatibility of generative AI with core tenets of arbitration. Analyzing just how and where the tensions arise, we find that the two are not such strange bedfellows after all.

III. We should start experimenting with generative AI

As of April 2023, apparently no one on the American Arbitration Association's roster was using ChatGPT. ((Fenwick, *supra* note 66.)) Or at least, no one admitted to using it. When faced with new technology, exercising some caution before jumping onto the bandwagon is commendable. Indeed ABA Resolution 112 urges the legal profession to address emerging ethical and legal issues concerning AI use, such as explainability and transparency, as well as the ethical and beneficial usage of AI. ((Accessible at: https://www.americanbar.org/content/dam/aba/directories/policy/annual-2019/112-annual-2019.pdf.)) Still, it is possible to be too cautious and miss the boat altogether. This part of the essay argues that the time is ripe to start experimenting with AI, considering its interaction with

trust and confidence in the arbitral process, as well as the importance of truth and of doing no harm.

A. Trust and confidence in arbitration.

"There was a time when arbitration was viewed disdainfully as an inferior process of justice. Those days are now well behind us." ((Tjong Very Sumito v. Antig Investments [2009] 4 S.L.R.(R). 732 (Sing.), [28].)) Having earned the trust of users, the arbitral community would do well not to squander it. The trust objection to using generative AI typically takes one of two forms: We cannot trust AI decision-making because it is a black box, ((Rebecca Wilson et al., *Busting the Black Box: Big Data Employment and Privacy*, 84 DEF. COUNSEL J, Jul. 2017, 6.)) and/or we cannot entrust our multi-billion dollar, bet-the-nation arbitration outcome to a mere AI.

It is true that AI is inexplicable, sometimes even to its creators, and that the use of robot judges (Estonia ((Tara Vasdani, "Estonia set to introduce 'AI judge' in small claims court to clear court backlog", THE LAWYER'S DAILY (Apr. 10, 2019), https://www.thelawyersdaily.ca/articles/11582/estonia-set-to-introduce-ai-judge-in-small-claims-court-to-clear-court-backlog-.)) and China ((Tara Vasdani, *Robot justice: China's use of Internet courts,* LEXISNEXIS CANADA (Feb. 2020), https://www.lexisnexis.ca/en-ca/ihc/2020-02/robot-justice-chinas-use-of-internet-courts.page.)) are early movers) has so far been confined to small claims cases. ((Zoppo, supra note 27.)) But if we accept the premise that tribunals should have the final say, it matters less that AI is producing the first draft.

Still, we might legitimately be concerned about the risk of arbitrators deferring too much to the AI-generated draft. Automation bias and automation complacency, referring respectively to the human tendency to unduly defer to or be over-confident in algorithmic reasoning and decisions, are real. ((Aileen Nielsen, *The Too Accurate Algorithm,* 9 CTR. L. & ECON. WORKING PAPER SER. 2022, 47.)) Accounts exist of individuals second-guessing what should be their informed judgments after cross-referencing their decisions against algorithmic output. ((*See generally* ch. 4, VIRGINIA EUBANKS, AUTOMATING INEQUALITY (2018) ("We all tend to defer to machines … [If] your research doesn't match the score, typically, there's something you're missing."))) Despite this, I suggest there is reason to believe the risks of this are manageable.

First (and I say this only half-jokingly), some lawyers of a certain stature are pretty self-important people who tend to believe they are right.

Second, the nature of legal decision-making. One reason for over-deference to algorithms is their seeming logic and infallibility in matters of calculation. But as Justice Holmes reminded us, the life of the law has not been logic. There is no single right answer like there is in math. As an example, consider how rule selection can involve discretion and the exercise of judgment. Karl Llewellyn famously compiled a list of 24 dueling canons of statutory construction ((Karl N. Llewellyn, *Remarks on the Theory of Appellate Decision and the Rules or Canons of About How Statutes are to be Construed,* 3 VAND. L. REV. 395 (1950).)) that illustrates the subjectivity of rule choice. "A statute cannot go beyond its text", yet "[t]o effect its purpose a statute may be

implemented beyond its text". ((*Id.*, 401)) Unambiguous language must be followed, unless a literal interpretation would lead to absurdity or mischief. ((*Id*., 403.)) Every word and clause must be given effect, unless inadvertently inserted or if repugnant to the statutory purpose. ((*Id*., 404.)) Leaving the realm of statutory interpretation we see the same duality in equitable maxims. In a variation on Newton's Third Law, for every maxim there is an equal and opposite rebuttal. In the adversarial system, counsel will need to figure out the rule that best advances their client's case and, depending on who you ask, the arbitrators' decision will be informed by their assessment of which choice of rule best accords with the law, their politics, or their sense of equity and justice.

Third, the collective decision-making process in multi-person tribunals serves as a bulwark against deferring too much to AI. It is not unthinkable for an individual all alone, staring at the magic of AI-generated input, to feel self-doubt. But the collective will of multiple people in conversation with each other is much harder to overbear. Discussions can also lead to disagreement and the sharpening of thought that comes from having to defend your view. While there is always the risk of groupthink and of blindly heeding the directions of a co-arbitrator (see gravitas, discussed at [5]), that is what tribunal diversity and proper selection of arbitrators are for.

All things considered, arbitrators steeped and socialized in legal culture will be inoculated enough against deference to an algorithm. Trust in the arbitral process and confidence, rooted in the belief that appointed arbitrators will make decisions based on independent exercise of judgment, will remain intact.

B. Truth

Legal decision-making, be it litigation or arbitration, is often described as a "search for truth". ((Dodson v. Persell, 390 So.2d 704, 707 (Fla. 1980).)) "The object of a lawsuit is to get at the truth and arrive at the right result." ((Marvin E. Frankel, *The Search for Truth: An Umpireal View,* 123 U. PENN. L. REV. 1031 ("Frankel"), 1035 (1975) (citing Justice David W. Peck).)) The focus on truth sits uncomfortably with generative AI at a theoretical level, since the AI's output is probabilistic. It does not "know" what it is saying. Only after the fact might we find that sometimes, the likeliest output happens to coincide with the truth. However, conceptual incongruity should not impede the use of generative AI in arbitration.

In some respects, we put too much store by the focus on explainability—understanding, and in turn understandability. Human judgment is not always explicable. Some theories of law ((Two theories are American legal realism, whose rule skepticism suggests that something other than legal rules decides legal cases—the ideology or "sense" of the decision-maker; and critical legal theory, which considers all law to be political, all the way down.)) posit that judicial reasoning works backwards from the desired end-result, which may in turn be arrived at by gut feel or a vague sense of where the equities of the case lie. Juries rarely, if ever, need to give reasons and we are fine with that. There was a time when parties were happy with non-speaking awards. And there is no coherent way to explain "aesthetic" judgments that are based on feelings or matters of personal taste. ((*Aesthetic Judgment*, STANFORD ENCYC. PHIL. (Feb. 13, 2023),

https://plato.stanford.edu/entries/aesthetic-judgment/.)) If the outcome works, are we tying ourselves unnecessarily into knots by insisting on questioning the process?

Further, we should beware over-romanticizing the pursuit of truth. Lofty ideals can be undercut by the advocate's loyalty to the client. ((Frankel, *supra* note 88, 1035–6.)) The cost of instituting proceedings sometimes means it is not the party with the best case who prevails, but the party with the stamina to outspend and outlast adversaries in a war of attrition. In trials, "truth" comes down to whether the burden of proof is met. ((*The Danger of Calling a Trial a "Search for the Truth"*, TEMPLE UNIV. BEASLEY SCH. L. (undated), https://law.temple.edu/aer/2017/05/18/danger-calling-trial-search-truth/.)) The impact of this is particularly stark in criminal cases, when the beyond reasonable doubt standard means accused persons can be factually guilty but legally innocent. Even in civil proceedings which arbitrations invariably are, what decision-makers strive for is not the truth but their best guess of it based on whatever evidence is available after time and memory take their toll. Considering the above, upholding truth is not incompatible with the use of generative AI.

C. Doing no harm (good citizenry)

Generative AI comes with the baggage of a whole panoply of problems, some of which have been referenced above. It can be biased, perpetuating discrimination. It can hallucinate, contributing to misinformation and disinformation. Its very genesis can be unethical, as some would say of OpenAI's use of Kenyan workers contracted on a pittance to sift through the gutters of the internet—violence, hate speech, sexual abuse—for the sake of training the AI. ((Billy Perrigo, *OpenAI Used Kenyan Workers on Less Than $2 Per Hour to Make ChatGPT Less Toxic*, TIME (Jan. 18, 2023), https://time.com/6247678/openai-chatgpt-kenya-workers/.)) We might wonder whether individuals and institutions become complicit or tainted by association by using this technology.

On the other hand, these problems are not peculiar to the use of AI in arbitration and are not the arbitral community's burden alone to bear. In this respect they are distinguishable from issues like the environmental impact of arbitration, concerns over which have rightly prompted course correction through protocols and model clauses for "greener" arbitration. ((Campaign for Greener Arbitrations, https://www.greenerarbitrations.com/.)) As far as generative AI is concerned, what is within our power to do is apply our judgment to AI output, lest its embedded ills leap from on-screen rhetoric to reality. And, as consumers choosing between competing service providers, to encourage or condemn through the persuasive power of the purse.

Conclusion

Generative AI, which will complement and not displace us, can be used compatibly with arbitration's cherished precepts. We should experiment to see how AI can improve legal practice. There will be a need to manage expectations and work out the rules of the road. The process calls for input from all stakeholders—institutions, arbitrators, counsel, clients—and will not be easy. But what successive innovations offer is a lens through which to reexamine the values, ideals, and direction of the arbitration community. Technological change makes all the old debates new once more. But we will come through again, as we did before.

Winner: California Arbitration 2023 International Arbitration & ADR Writing Competition

---