CHAPTER 3

The New York Convention's Borderline. Blockchain Arbitration and Artificial Intelligence

MATTHIAS LEHMANN[1]

New technologies are disrupting the world of dispute resolution. Blockchain arbitration (BA) allows a community of users across the globe to act as arbitrators.[2] Artificial intelligence (AI) even offers the prospect of replacing human arbitrators with algorithms.[3] Are these new phenomena compatible with the New York Convention (NYC)?[4] In particular, must states recognise and enforce decisions rendered via BA or AI under the Convention?

Some may deem this discussion moot. After all, the new technologies harbour the possibility of 'self-enforcement', for instance via a smart contract that holds digital assets of the defendant in escrow and automatically transfers them to the winning claimant.[5] However, not all technological dispute resolution methods feature a self-enforcement mechanism. A reason may be that defendants are reluctant or unable to put their digital assets in escrow. And even where they do, the amount may not suffice to satisfy the award.

Therefore, award enforcement by real people in the real world remains important. To this end, the compatibility of the new dispute resolution methods with the NYC is relevant. One can hardly overstate the significance of this question. If it were to be answered affirmatively, this could herald a true modernisation of arbitration. We would witness a pivotal moment in time, in which technology upheaves dispute resolution. No longer would it be necessary for humans to argue and converse with each other over the proper solution of the case. Instead, entirely different ways of solving disputes would be opened. But if the question were to be answered negatively, it would mean that the traditional way of solving disputes would endure in its importance. Human arbitrators would still have to interact with each other and debate over the right decision in a case. This could change only if a new, technology-open convention were elaborated and ratified by at least a similar number of states.

DOI: 10.54195/FMVV7173_CH03

1 Blockchain Arbitration

1.1 Description and Differences to Traditional Arbitration

Blockchain Arbitration is used here to describe arbitration procedures conducted on the blockchain, in accordance with the logic and rules inherent to the blockchain environment. This definition, therefore, differs from the definition of BDR of blockchain dispute resolution used in Chapter 2.[6] A typical BA is quite different from ordinary arbitration. As a point of reference, I am using Kleros, which is currently the leading system in this area.[7] Kleros offers a way of solving disputes that is entirely digital. To start with, the notice of arbitration and the statement of claim will be sent electronically.[8] The claimant will typically have to provide security for the fees by 'staking' crypto assets, which can be compared to putting them into a digital vault.[9] 'Jurors' – i.e. decision-makers – will be selected algorithmically, typically from a roster of the members of a Decentralized Autonomous Organization (DAO); they will also have to stake crypto assets to show their seriousness and willingness to act as jurors.[10] The claim, the defence and all pieces of evidence will be sent electronically to the selected jurors. They will also render their decision electronically in cyber-space, without knowing each other or meeting physically or virtually. Even the arbitration agreement may be made digitally, for instance by ticking a box in a window that pops up when the contract is concluded on the internet.

The most crucial difference to traditional modes of dispute resolution, however, is the way in which the jurors will be rewarded. Those who voted in accordance with the majority will receive a fee and additional crypto assets on top of those that they staked. In contrast, those in the minority will not receive any fee; in addition, they will see their staked crypto assets either fully or partially taken away. (As a matter of fact, those crypto assets will serve to reward the jurors that are in the majority.) The purpose of this system is to incentivise jurors to vote honestly.[11] Given that they act anonymously, they would lose nothing by voting in bad faith. An incentive scheme has therefore been designed to ensure the proper decision is reached.

The fundamental idea underpinning BA is that a majority of the people will arrive at the right decision. In other words, BA relies on the 'wisdom of the crowds'.[12] The linchpin of this idea is the so-called Schelling point, after the Nobel prize-winning game theorist Thomas Schelling. Schelling illustrated his concept by an experiment in which participants were asked to meet in New York City the next day, but were not given the exact time and place.[13] Although they could not communicate with one another, a majority chose to come at noon to the information booth at Grand Central Station. This was meant to show that a majority of reasonable people, even when deciding in absence and not being able to communicate with each other, will agree on a certain result, which Schelling calls the 'focal point'. The focal point can be understood as proving the 'wisdom of crowds', which supposedly exceeds that of any single individual.

Some systems allow an appeal against any decision rendered.[14] In order to make full use of the wisdom of the crowds, each appeal will be decided by double the number of jurors as before. At the same time, the fees also increase. While the appeal can be repeated infinitely, the increase in the fees is designed to incentivise parties to abstain from unnecessary appeals.

1.2 The Requirements of the NYC

It is doubtful whether this model is sufficient to qualify BA processes as 'arbitration' under the New York Convention. As a matter of fact, many of the features of the new type of dispute resolution are at odds with the Convention's provisions.

For starters, Art II(1) NYC requires an agreement 'in writing', and it is by no means sure that this also covers an electronic exchange of communication, let alone agreements made on a blockchain.[15] Second, Art V(1)(b) NYC requires that each party is able to present its case, which it may be inhibited to do in a blockchain context, where the parties do not interact with the jurors. Third, Art I(1) NYC starts from the idea that the award is made within the territory of a certain state, however such a 'seat state' is impossible to locate in the decentralised environment of BA. Fourth, Art IV(1)(a) NYC calls for a 'duly authenticated original' of the award, which may be difficult to produce where jurors vote in cyber-space. Fifth, the typical

motivation of awards is highly unusual in BA, where jurors cast their votes independently of each other and most often need not give any reason why they voted in a certain way.

Still, none of these problems is unsurmountable.[16] One may for instance interpret the 'in writing' requirement broadly in line with the principle of equivalence of written and electronic form;[17] the parties in BA can at least submit their pleadings and any evidence to the jurors; the seat can be fixed by agreement in a particular country; the award could be signed electronically by the jurors; and the parties could be deemed to have renounced the requirement of a motivated award, as it is also possible in decisions made *ex aequo et bono*. This shows that the NYC requirements, if taken in isolation, may be complied with by BA.

1.3 The Notion of 'Arbitration'

The most fundamental problem is, however, whether BA may be at all considered as 'arbitration' in the sense of the NYC. Though the letter of each of its provisions can be complied with, BA could be completely outside its scope. That is because BA, contrary to what its name suggests, may not be arbitration, but something entirely different.[18]

Whether this is the case or not is particularly hard to answer given the lack of a definition of arbitration in the Convention itself. There is no universal agreement of what arbitration is either.[19] The well-known debate between those likening it to a contract and those stressing its similarity to judicial proceedings does not need to be repeated here.[20] As is also well-known, the first view tends to result in a greater role of party autonomy than the second. Proponents of contractual characterisation will therefore most probably consider BA as arbitration provided the parties have agreed on it, while those stressing the quasi-judicial function of arbitration might lean to the opposite conclusion.

Independently of the 'true' nature of arbitration – which may not be summed up with any one category but is more likely to be hybrid – there are important arguments to allow new technologies in arbitration. In particular, the right of the parties to fashion the proceedings according to

their own preferences is almost universally recognised.[21] If they wish to select the jurors randomly through an algorithm and have them vote anonymously and separately, why stop them from doing so? There is no doubt that BA could harness the advantages of new technological developments by broadening the roster of arbitrators and streamlining the procedure. As a matter of fact, this could offer a cost-effective way of solving disputes, especially those of smaller monetary value.

On the other hand, it is hard to deny a certain tension between BA and the notion of arbitration as it is traditionally understood. Even those authors who are stressing the contractual nature of arbitration admit that it also has an adjudicative function.[22] This is not merely a theoretical debate, but relevant for the scope of the NYC. Though it does not define the notion of arbitration, one can see through its scarce provisions and pithily drafted words a particular image of arbitration, which the drafters must have had in mind. It speaks of the 'appointment of the arbitrator', as if it was done regularly by the parties themselves and only rarely by an institution. It also mentions the 'arbitration proceedings', which alludes to a certain procedure modelled on or at least vaguely similar to judicial proceedings. This type of procedure offers parties the opportunity 'to present their case', to have some interaction with the court and respectively the arbitral tribunal, to engage in a rational debate. Importantly, these are also minimum requirements under international treaties for access to justice.[23]

Still, none of this proves that the notion of 'arbitration' could not also be fulfilled by algorithmically chosen arbitrators or jurors. Decisive are, however, the effects of arbitration. As clarified by Art II(1) NYC, a (valid) arbitration agreement excludes the jurisdiction of state courts. And as set out in Art III(1) NYC, contracting states must recognise the award and enforce it in accordance with the rules in place in their territory. This means indeed that arbitration is more than a pure contract. It discards the jurisdiction of state tribunals, and it produces awards with *res judicata* effect. To achieve this result, certain requirements are necessary. In particular, the arbitrators must render a decision they believe to be fair and just, and in accordance with the rule of law. They must be unbiased and hence cannot remain anonymous. And they must, during the course of the arbitral proceedings, converse with each other in order to find the proper solution to the case.

For these reasons, not everything that the parties agree on deserves to be called 'arbitration', as is known from the lengthy discussion on the delineation from expert determination and other phenomena.[24] It seems obvious that if the parties were agreeing to solve their dispute by the toss of a coin, the coin toss would not be considered an award. But nor would it be if a private person were to witness it and put the result into the form of a written award. Parties may authorise arbitrators to disregard the law and decide *ex aequo et bono*.[25] But they cannot authorise them to decide in accordance with a standard beyond both law and equity. This would no longer be adjudication because adjudication is not merely about deciding a dispute. An essential ingredient and characteristic of any form of adjudication is the search for a decision that is just or fair.[26] As will be shown next, BA is not designed to achieve such a result.

1.4 The Discrepancy of BA

The scheme underpinning BA, in particular the Schelling model, cannot be compared to traditional models of arbitration and the concept underlying the NYC. That is because it does not incentivise jurors to search for a just and fair decision. It incentivises them to think about what others think is a just and fair decision. The Kleros Whitepaper is very open about this point by admitting that through its system,

> *"... parties are incentivized to vote what they think, other parties think, other parties think. . . is honest and fair".*[27]

But whether you think something is honest and fair, or whether you shall estimate if others may deem it to be honest and fair, are two different things entirely.

An example will illustrate why this discrepancy matters. Suppose three jurors (A, B and C) are informed about a criminal accusation and are asked whether the defendant is guilty. Suppose further A and B to be highly intelligent and aware of their rare gift. Both conclude on the defendant's innocence, while the averagely intelligent juror C is convinced of the defendant's guilt. In any system of adjudication, A and B should be incentivised to vote according to their true opinion. However, in the Schelling system of conflict resolution, it is different. In this system, A and B have to

think about what the majority will decide. They are incentivised to vote with that imagined majority, even if they are convinced the truth was different. Not knowing each other and not being able to communicate with each other, A and B will conclude an ordinary juror is likely to find the defendant guilty. Rationally, they will vote this way, in order to gain a fee and to avoid losing their stake.

The following table illustrates the different voting behaviour:

	SYSTEM OF ADJUDICATION	BLOCKCHAIN ARBITRATION
A	not guilty	guilty
B	not guilty	guilty
C	guilty	guilty
OUTCOME	not guilty	guilty

This difference must not surprise us. It reflects the very old insight that the majority is not always right. Where jurors can deliberate with other jurors, they are able to argue, show them their mistakes, and convince them of their opinion. This is an important safeguard for any form of adjudication. A forceful example is shown in the famous movie "Twelve Angry Men", in which a single juror disagrees with a guilty verdict, holds out and finally manages to convince the others of the accused's innocence.[28] Scores of other incidents have also proven the value of jury deliberations.

In contrast, jurors in BA cannot communicate with each other and are in addition penalised for having a different opinion. As a consequence, they will not make their voices and doubts heard. Instead, they are most likely to engage in herd behaviour, voting like the majority is likely to vote. Worse still, they will not even vote like they believe the majority is voting, but what they think the majority of jurors is likely to think the majority thinks. Consider again the Kleros Whitepaper:

> *"... parties are incentivized to vote what they think, other parties think, other parties think. . . is honest and fair."*[29]

This amounts to a bet on the probable outcome of a vote, not the vote itself. Such bets follow a very different logic than giving one's own opinion. They will vote not in accordance with their opinion of what is fair or

just, but in statistical or stochastic terms. The result of such an exercise is only a faint shadow of adjudication according to the proper convictions of the jurors, as it was envisaged by the NYC.

1.4 Wager as Adjudication?

The question, in dry legal terms, is whether BA can be assimilated to arbitration regardless of these flaws. After all, the NYC does not prescribe a particular type of procedure. Instead, it leaves much of this to the parties and to the laws of contracting states. Also, the image underpinning the NYC is not cast in stone but may be subject to change.

An example is baseball arbitration, in which arbitrators are not free to decide as they prefer but are bound to choose between either the full petition made by the claimant or the full petition made by the defendant.[30] Still, there is common agreement that this form of arbitration qualifies for award recognition and enforcement under the NYC.[31] There are at least two crucial differences between baseball arbitration and BA. First, the incentive structure: baseball arbitrators' remuneration is not conditioned on their hitting the 'right' opinion, but independent of it. Second, the outer limit of the decision is not set by the opinion of their co-jurors, but by the claims of the parties. In this sense, baseball arbitration is just common arbitration 'with a twist'.

It is apparent that the contracting states of the NYC did have a particular concept of 'arbitration' in mind when signing the Convention in 1958. This concept was that of one or several persons' search for a just and fair decision. While this does not result from any single provision of the Convention, it clearly follows from all of them read together. Fair and just results cannot only be achieved by proceedings similar to those in a courthouse. Nevertheless, some basic procedural conditions must be met to at least enable a fair and just result (although this will not always be the result). Chief among them is – in the case of several jurors – a possibility for deliberation between the jury members about the right decision, and a proper incentive scheme.

BA – at least in its current form – does not live up to this concept. It is essentially a betting system, similar to event prediction mechanisms. It is

correct that one cannot expect arbitrators or jurors to be right 100% of the time.[32] Yet BA is designed in such a way that the right result is not even looked for, just the result that the majority will probably think is right. It is hard to imagine that the states of the NYC signed up to enforce the results of such a mechanism and were willing to renounce their right to exercise their own judicial power in its favour. Gaming cannot replace adjudication.

2 Artificial Intelligence

2.1 Large Language Models as the Test Case

Within the last few years, AI has developed spectacularly. This makes the replacement of human arbitrators by computers and algorithms a distinct possibility. There are many different variants of AI.[33]

Particularly impressive are the achievements of so-called Large Language Models (LLMs), the most prominent of which is ChatGPT.[34] LLMs are very good at summarizing text, answering all types of questions, and generating text. They are based on previous learning from massive amounts of pre-existing data, which they take from sources like the internet.[35] These data inputs are put through a so-called 'transformer'.[36] It consists of an encoder, which transforms patterns of words and sub-words into numerical tokens as input, and a decoder, which produces an output sequence, along with several other, hidden layers. These layers together form a complex architecture, which is designed to mimic the human nerve system (therefore also called 'neural network').[37] It can work on its own without any human supervision or guidance.[38] In contrast to previously dominant sequential methods, a transformer does not process one word at a time, but whole sentences, paying attention to the so-called 'embeddings', or specific context in which a word or sub-word is used.[39] This helps to achieve better quality results ('deep learning') while also speeding up the process of machine learning because several meanings and their interactions can be learned in parallel. The output sequence is generated from this pre-trained material (therefore the name 'Generative Pre-trained Transformer' – GPT). It is a prediction of the word or other output that will most probably appear after a series of other words.

LLMs are used in all sorts of areas of life, whether it is business, education, research, or just for fun.[40] They may also find their way into dispute resolution. Some projects have already been developed to make use of them in the field of mediation.[41] LLMs may also become significant for arbitration. Already now, they are employed for a variety of purposes, such as note taking, transcription, factual and legal research, or as assisting in the drafting of briefs, pleadings, and awards.[42] It is not beyond imagination that one day they might be honed to decide the dispute themselves.

2.2 Problems of AI

Yet AI models and LLMs in particular are beset by a plethora of problems.[43] For one, the quality of the results is highly dependent on the quality and depth of the dataset ('garbage in, garbage out').[44] Moreover, the variations of the syntax of the question put to the LLM (the so-called prompt) may lead to dramatic changes in output.[45] ChatGPT may even be biased towards specific genders, races, ethnicities, or other social groups.[46] Sometimes, LLMs simply produce wrong information, which may seem deceptively accurate (so-called hallucinations).[47]

Limitations like these have led the literature to warn against using AI for arbitration.[48] It was specifically fretted that an algorithm trained on precedent may be infected with human biases and even exaggerate them by holding them as 'true' for its future decisions or outcome predictions.[49] But computers do not suffer from human fallacies like the 'lunch break bias'.[50] They render predictions with stochastic exactness. If anything, they will prune out biased human decisions as outliers and focus on the majority of the case law. This makes it highly unlikely that parties will be treated worse than before a human arbitrator.

Another problem of AI is its incapacity of rendering properly motivated decisions. Although it can write text, this is not comparable to a reasoned award since it does not *explain* the real reasons why a certain decision was rendered.[51] As these programmes learn without supervision on their own, even their coders struggle to explain the results they produce.[52] Indeed, while algorithms can give reasons for their results, these are not comparable to human reasons in the sense of the *beliefs* of a human judge or arbitrator. Yet everyone who decries LLMs for the lack of giving true reasons

should take a critical look in the mirror: Are human judges and arbitrators always giving the *real* reasons for the decisions they take? Or are they, more often than not, using formal legal arguments, which hide their true motivations? One must neither be a legal realist nor a cynic to believe the latter. Any form of legal decision will undergo a complex way of formalisation before it is reached, with a view to make it acceptable and convincing for all members of the tribunal, for the parties of the proceedings, and potentially for third parties such as the arbitral institution. This formal rationale will mostly mask the previous dispositions and personal beliefs of the arbitrators as well as the discussion between the members of the tribunal and the different ideas underpinning the decision. By comparison, computers are more transparent than humans because their memory is not an internal thought process but hidden in plain sight on a computer chip.

2.3 Facing the NYC

But could decisions rendered by AI be considered awards in the sense of the NYC? Again, a dry look at the provisions is crucial to answer this question. Quite soon, doubts will appear. The Convention's concept of an award is a decision rendered either by arbitrators or by a permanent arbitral tribunal (see Art I(2) NYC). Even in the latter hypothesis, an arbitrator must be appointed, and the parties notified of it (see Art V(1)(b) NYC). Notions like 'arbitration proceedings' and the 'arbitral procedure' seem to imply human interaction, as these terms have so far been understood as referring to written or oral presentations, arguments and debates between human beings. An algorithm will also have difficulties to produce a 'duly authenticated award' (Art IV(1)(a)) if 'authentic' is understood as being attributable to a specific person.

Yet these are only superficial reasons for rejecting an AI decision's recognition and enforcement under the Convention. While each of them carries some weight, none of them fully explains why such a decision cannot be considered an 'award' and why the process leading up to it cannot qualify as 'arbitral proceedings'. The most important reason lies in the technology itself. AI, at least in its present form, does not work like human intelligence of a judge or arbitrator. It does not search for a fair or just decision. Instead, it is based on statistical probabilities. Models like LLMs are nothing else than 'stochastic parrots'.[53] Their function is to predict the next

word or sentence from prompts. In this sense, they resemble BA, the difference being that the prediction is done by algorithms and not by humans. If one refuses to consider BA as arbitration, then one must reject AI as arbitration even more emphatically. LLMs merely produce a prophecy of what human arbitrators might do based on what they have been doing in past cases. But they do not provide an opinion of what is fair and just in the specific circumstances of the new case at hand. They are at best a vague image of justice.

That does not exclude the use of AI in arbitral proceedings. It can fulfil many useful purposes, from analysing records to organising documents submitted to the tribunal and translating foreign languages. They may also be used by arbitrators in the preparation of the draft award. However, a human being must verify the draft and, crucially, form an opinion on the proper solution of the dispute. In other words: while AI may be used for assistance, it cannot replace the reasoning of the human arbitrator(s). Anything else would be incompatible with the notion of arbitration underlying the NYC. At least current models of AI, the LLMs, do not fulfil this requirement.

2.4 Workarounds

Since BA and AI dispute resolution do not qualify as arbitration, the question arises whether it is nevertheless possible to have their products recognised and enforced under the NYC. Ways of doing so could be imagined. One way of doing so would be to put a settlement of the parties before an arbitrator, who puts it into an award. Another would be for one or several human arbitrators to just adopt a decision rendered by BA or an AI system.

There is already one practical example for the latter approach.[54] In 2020, a dispute between two Mexican parties over a rental contract was submitted to the BA system Kleros. This system rendered a decision via three jurors. On 27 November 2020, a (sole) human arbitrator in Mexico then copied the BA decision into an 'award'. While the arbitrator pretended that he reached a decision himself, it was revealed that he simply copied the BA decision. The winning party filed a motion for recognition and enforcement of the award. On 28 May 2021, the District Court of Jalisco granted this motion.

Proponents of BA and AI arbitration will cheer over this 'creative' solution. When looked at superficially, it complies with the requirements of the NYC. The procedure was in accordance with the wishes of the parties. A human being is certifying an award. Effectively, the arbitrator 'decides' that the dispute has already been resolved by the BA or AI mechanism. What could possibly be wrong with that?

The arbitrator who renders an award in this way does not actually decide the dispute. She merely certifies a result that has been reached by other means. This is not arbitration – after all no human arbitral tribunal has searched for a fair and just solution. It does not comply with the NYC, but merely circumvents its requirements. It is an avoidance of the law, a *fraus legis*, which must not be permitted. The fact that an arbitrator signs off on an award does not make the content and the genesis of the award meaningless. The latter remains essential for assessing the nature of the decision rendered.

In a similar vein, a decision rendered by an arbitrator on a BA or AI decision partakes in the legal nature of the latter. It is not an 'award' in the sense of the NYC and does not enjoy the privileges under this Convention. If it were otherwise, anything could be certified as an 'award' and create *res judicata* effect in other signatory countries. While it is true that there are settlements that can circulate as awards, these are always preceded by proper arbitral proceedings. Settlements achieved without such proceedings are not considered as 'awards'. So it is for BA and AI decisions as well.

Yet it is possible that the arbitrator or the arbitral tribunal uses AI in the arbitral proceedings, for instance to administer documents, to gather information, or even to draft the award.[55] Is that impermissible? This is of course a can of worms, which cannot be opened and killed one by one here. But some guiding principles can hopefully be agreed upon: First, AI can be a useful tool to facilitate many of the arbitrators' tasks, especially when large amounts of information have to be processed, provided the obligation of confidentiality is fully complied with. Second, a complete prohibition of AI would be futile as it could hardly be monitored. Third, there are no pervasive arguments against employing technology as long as a human being is supervising its use and is accountable for the result. The last point

is perhaps the most important. It is this point that sets any form of AI assistance of human arbitrators apart from AI arbitration.

3 Conclusion

Current models of BA and AI are incapable of producing an enforceable award under the NYC. They may be improved in the future. It is by now unforeseeable which other models will be invented. Whether they will satisfy the Convention's requirements cannot be decided at this point. One can only assess those models that presently exist, not those still to be developed.

It is possible, however, to define some key requirements. The most essential for any 'award' in conformity with the conception of the NYC will be a human arbitrator voicing his or her own personal opinion about the fair and just resolution of the case. In the case of an arbitral tribunal, a further key requirement will be that the arbitrators have the opportunity to exchange arguments with each other. Unless these requirements are met, the resulting decision does not qualify as an 'award' to be recognised and enforced under the Convention.

That must not be the end of the story for technology in dispute resolution. Quite to the contrary, technology such as AI can be a helpful tool to assist the arbitral tribunal. However, pure BA or AI arbitration are incompatible with the NYC and their products do not benefit from its recognition and enforcement rules. If the states feel the political will to change this, they need to either modify the Convention or conclude another one.

Notes

1 I wish to thank Paul Lorenz Eichmüller and Laurenz Faber for helpful comments and research assistance.

2 On blockchain arbitration in general, *See* Chevalier, Maxime. "From Smart Contract Litigation to Blockchain Arbitration, a New Decentralized Approach Leading Towards the Blockchain Arbitral Order', (2021) 12 *Journal of International Dispute Settlement* 558-584 (hereafter Chevalier); Hourani, Sara. "The Resolution of B2B Disputes in Blockchain-Based Arbitration. A Solution for Improving the Parties' Right of Access to Justice in the Digital Age?" de

Oliveira, Leonardo & Hourani, Sara. *Access to Justice in Arbitration: Concept, Context and Practice* (2020) 251-272; Kadioglu Kumtepe, Cemre. "A Brief Introduction to Blockchain Dispute Resolution" (2021) 14 *John Marshall Law Journal* 138-157; Kindt, Torsten. "Blockchainbasierte dezentrale Streitbeilegungsverfahren und ihr Verhältnis zur Schiedsgerichtsbarkeit" (2023) 21 *SchiedsVZ/German Arbitration Journal* 241-254 (hereafter Kindt); Metzger, James. "Decentralized Justice in the Era of Blockchain" (2018) 5 *International Journal on Online Dispute Resolution* 69-81; Ortolani, Pietro. "Blockchain Technology and Arbitration" Ortolani, Pietro and others, *International Arbitration and Technology* (Wolters Kluwer 2022) 239-252; Raneda, Julie. "Decentralized Justice Systems, a New Player in the Field of Alternative Dispute Resolution?" (2023) 21 *ICCA Congress Series* 695-706; Wiegandt, Dirk. "Blockchain and Smart Contracts and the Role of Arbitration" (2022) 39 *Journal of International Arbitration* 671-690 (hereafter Wiegandt); Zh UK, Alesia. "Applying Blockchain to the Modern Legal System: Kleros as a Decentralised Dispute Resolution System" (2023) 4 *International Cybersecurity Law Review* 351-364.

3 Chan, Elizabeth, Nasir Gore, Kiran & Jiang, Eliza. "Harnessing Artificial Intelligence in International Arbitration Practice" (2023) 16 *Contemporary Asia Arbitration Journal* 263-295; Scherer, Maxi. "Artificial Intelligence and Legal Decision-Making: The Wide Open?" (2019) 36 *Journal of International Arbitration* 539-574 (hereafter Scherer).

4 On this question, *See* e.g. Boronkay, Miklós & Exenberger, Philip. "Blockchain, Smart Contracts and Arbitration, Overrated Hype or Chance for the Arbitration Community?" (2020) *Austrian Yearbook on International Arbitration* 411-418 (hereafter Boronkay & Exenberger); Sanyal, Arijit. "Can The New York Convention Stand the Test of Technology Posed by Metaverse Awards?" *Kluwer Arbitration Blog* (20 December 2022); Sharma, Chandrika. "Blockchain Arbitral Award: Potential Challenges in Recognition and Enforcement under the New York Convention" (2022) 16 *Revista Romana de Arbitraj* 85-107; Wilinski, Piotr. "Should the Miami Draft Be Given a Second Chance? The New York Convention 2.0" (2019) 34 *Spain Arbitration Review/Revista del Club Español del Arbitraje* 77-90; Yang, Yueh-Ping. "The Crowd's Wisdom in Smart Contract Dispute Resolution: Is Crowdsourced Dispute Resolution Arbitration?" (2022) 15 *Contemporary Asia Arbitration Journal* 175-204 (hereafter Yang).

5 *See* Chevalier 583 (highlighting that blockchain arbitration enjoys self-enforcement capabilities); Wiegandt 686 (noting that allowing arbitrators to implement decisions directly on-chain into the smart contract's code).

6 *See* p. 40. *See* also the definition in Chapter 7.

7 *See* the Kleros whitepaper (Lesaege, Clément, Ast, Federico & George, William. "Kleros Short Paper v1.0.7"). https://kleros.io/whitepaper.pdf

8 *See* the example in Kleros whitepaper p. 6.

9 Kleros whitepaper p. 6-7; Wiegandt 682.

10 Kleros whitepaper p. 4; Kindt 244; Wiegandt 682.

11 Kleros whitepaper p. 4.

12 The idea of the 'wisdom of the crowds', while going back to Antiquity, was popularised by Surowiecki, James. *The Wisdom of Crowds: Why the Many are Smarter than the Few and How Collective Wisdom Shapes Business, Economics,*

Societies and Nations (Doubleday 2004). It is today often used in the context of blockchain-based dispute resolution, *See* e.g. Yang (n 3) 175.

13 Schelling, Thomas C. *The Strategy of Conflict* (2d ed., Harvard University Press 1981), p. 55.

14 Kleros whitepaper p. 7.

15 On this question, *See* e.g. Boronkay & Exenberger 416-417; Kindt 248.

16 Chevalier 568-573 (showing how the seeming incompatibility of BA with traditional arbitration features can be overcome).

17 UNCITRAL, Recommendation regarding the interpretation of article II, paragraph 2, and article VII, paragraph 1, of the Convention on the Recognition and Enforcement of Foreign Arbitral Awards, done in New York, 10 June 1958, adopted by the United Nations Commission on International Trade Law on 7 July 2006 at its thirty-ninth session, 2006: A/6/17. The Recommendation refers, inter alia, to the United Nations Convention on the Use of Electronic Communications in International Contracts, done on the 23 November 2005 in New York, Art 9(2) of which prescribes that electronic communication meets the requirements of a written form provided 'the information contained therein is accessible so as to be usable for subsequent reference'.

18 *See* in this regard the position of Cemre Kumtepe Kadıoğlu in Chapter 2, who takes the position that crowdsourced voting systems are not corresponding to the definition of arbitration.

19 Yang 187-188 (underlining that the definition of arbitration remains unclear).

20 For a summary of this debate, *See* Born, Gary B. *International Commercial Arbitration* (3d ed. (update Nov 2023), Wolters Kluwer 2023) § 1.05[A]–[B]; Robert, Jean. *L'arbitrage: droit interne, droit international privé* (6th ed., Éditions Dalloz 1993) pp. 3-10 (hereafter Born).

21 *See*, for example, Art 19(1) UNCITRAL Model Law on International Commercial Arbitration ("the parties are free to agree on the procedure to be followed by the arbitral tribunal in conducting the proceedings"). *See* also Blackaby, Nigel and others. *Redfern and Hunter on International Arbitration* (7th ed., Oxford University Press 2023) para. 6.07; Borris, Christian and others. Article V, in Wolff, Reinmar (ed.), *New York Convention: Article-by-Article Commentary* (2d ed., C.H. Beck 2019) paras 308-311 (hereafter Wolff).

22 Born (n 17) § 1.05[B]; Robert (n 17) 9-10; Ortscheidt, Jérôme & Seraglini, Christophe. *Droit de l'arbitrage interne et international* (1st ed., Montchrestien 2013), pp. 20-23.

23 *See* Art 6 European Convention on Human Rights.

24 *See* on this debate, *e.g.* Born § 2.02[C][2][B]; Tustin, Marcin. "Enforcing an Expert Determination Award Under the New York Convention" (2015) 28 *New York International Law Review* 25-46; Valasek, Martin & Wilson, Frédéric. "Distinguishing Expert Determination from Arbitration: The Canadian Approach in a Comparative Perspective" (2013) 29 *Arbitration International* 63-87; Ehle, Bernd. Article V, in Wolff paras 32-35. For the proposal of a convention for the recognition of expert determination agreements and decisions Saidov, Djakhongir. "An International Convention on Expert Determination and Dispute Bords" (2022) 71 *International & Comparative Law Quarterly* 697-726.

25 *See* Art 28(3) UNCITRAL Model Law.

26 *See* Born § 1.05[B] ("arbitration agreement does not produce a typical 'commercial' bargain, but instead results in a particular kind of dispute resolution process").
27 Kleros whitepaper, p. 8 (three dots in the original).
28 *See* https://www.imdb.com/title/tt0050083/.
29 *See* again Kleros whitepaper, p. 8 (emphasis mine).
30 For a description of baseball arbitration and its sub-forms, *See* Borris, Christian. "Final Offer Arbitration from a Civil Law Perspective" (2007) 24 *Journal of International Arbitration* 307, 308-311.
31 Mitrovic, Laurence. "L'arbitrage baseball: arbitrage ou mode alternatif de reglement?" (2003) 4 *Revue de l'Arbitrage* 1167, 1192. *See* also Borris (n 27) 317 (for Germany and other civil law jurisdictions); Timothy, Gracious & Duggal, Kabir. "Mixed-Mode Dispute Resolution" (2021) 11 *Conciliation and Mediation in India* 331, 349 (for India). Expressing doubts, however, Born (update Aug 2022) § 2.02[C][2][h].
32 *See* Kleros whitepaper, p. 8 ("We cannot expect jurors to be right 100% of the time. No arbitration procedure could ever achieve that.").
33 For a general overview, see Surden, Harry. "Artificial Intelligence and Law: An Overview" (2019) 35 *Georgia State University Law Review* 1305, 1310-1321.
34 On LLMs in general, see Myers, Devon and others. "Foundation and Large Language Models: Fundamentals, Challenges, Opportunities, and Social Impacts" (2023) *Cluster Computing* https://doi.org/10.1007/s10586-023-04203-7.
35 On the relation between the amount of the documents on which AI has been trained and the quality of its output, see Kandpal, Nikhil and others. "Large Language Models Struggle to Learn Long-Tail Knowledge" *Proceedings of the 40th International Conference on Machine Learning* (PMLR 2023). https://proceedings.mlr.press/v202/kandpal23a.html.
36 The foundational paper on the transformer is Vaswani, Ashish and others. "Attention Is All You Need" *Advances in Neural Information Processing Systems* (NIPS 2017) (hereafter Vaswani & ors.) https://proceedings.neurips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html."plainCitation":"The foundational paper on the transformer is Ashish Vaswani and others, 'Attention Is All You Need', Advances in Neural Information Processing Systems (Curran Associates, Inc 2017
37 Min. Bonan and others. "Recent Advances in Natural Language Processing via Large Pre-Trained Language Models: A Survey" (2024) 56 *ACM Computing Surveys* 1. On the different types of architectures for LLMs, see Naveed, Humza and others. "A Comprehensive Overview of Large Language Models" https://arxiv.org/pdf/2307.06435.pdf at II.J.; Wang, Thomas and others. "What Language Model Architecture and Pretraining Objective Work Best for Zero-Shot Generalization?" P*roceedings of the 39th International Conference on Machine Learning* (PMLR 2022). https://proceedings.mlr.press/v162/wang22u/wang22u.pdf at 2.1.
38 Radford, Alec and others. "Language Models Are Unsupervised Multitask Learners" https://d4mucfpksywv.cloudfront.net/better-language-models/language-models.pdf.

39 Vaswani & ors"plainCitation":"The foundational paper on this is Ashish Vaswani and others, 'Attention Is All You Need', Advances in Neural Information Processing Systems (Curran Associates, Inc 2017.

40 On the various applications, see e.g. Kocoń, Jan and others. "ChatGPT: Jack of All Trades, Master of None" (2023) 99 *Information Fusion* 101861; Myers and others (n 31).Master of None\\uc0\\u8217{} (2023

41 See Westermann, Hannes, Savelka, Jaromir & Benyekhlef, Karim. "LLMediator: GPT-4 Assisted Online Dispute Resolution" *CEUR Workshop Proceedings* https://ceur-ws.org/Vol-3435/paper1.pdf; Doernhoefer, Gary. "Artificial Intelligence and Dispute Resolution: A Primer and Opportunity for the Future" *ADR Notable* (21 March 2023). https://www.adrnotable.com/ai-primer/.

42 Chan, Elizabeth, Nasir Gore, Kiran & Jiang, Eliza. Harnessing "Artificial Intelligence in International Arbitration Practice" (2023) 16 *Contemporary Asia Arbitration Journal* 263, 274-281.

43 See e.g. Kaddour, Jean and others. "Challenges and Applications of Large Language Models" *arXiv* (19 July 2023) (hereafter Kaddour & ors.) http://arxiv.org/abs/2307.10169.

44 On the relation of the quality of the results with the number of documents fed into an LLM, see Kandpal and others (n 32).

45 Kaddour & ors. 27-28 (speaking of 'brittle evaluations').

46 Ferrara, Emilio. "Should ChatGPT Be Biased? Challenges and Risks of Bias in Large Language Models" (2023) *First Monday* http://arxiv.org/abs/2304.03738."plainCitation":"Emilio Ferrara, 'Should ChatGPT Be Biased? Challenges and Risks of Bias in Large Language Models' [2023] First Monday http://arxiv.org/abs/2304.03738 accessed 9 December 2023.","noteIndex":47},"citationItems":[{"id":3182,"uris":["http://zotero.org/users/2405672/items/AE6BK3VA"],"itemData":{"id":3182,"type":"article-journal","abstract":"As the capabilities of generative language models continue to advance, the implications of biases ingrained within these models have garnered increasing attention from researchers, practitioners, and the broader public. This article investigates the challenges and risks associated with biases in large-scale language models like ChatGPT. We discuss the origins of biases, stemming from, among others, the nature of training data, model specifications, algorithmic constraints, product design, and policy decisions. We explore the ethical concerns arising from the unintended consequences of biased model outputs. We further analyze the potential opportunities to mitigate biases, the inevitability of some biases, and the implications of deploying these models in various applications, such as virtual assistants, content generation, and chatbots. Finally, we review the current approaches to identify, quantify, and mitigate biases in language models, emphasizing the need for a multi-disciplinary, collaborative effort to develop more equitable, transparent, and responsible AI systems. This article aims to stimulate a thoughtful dialogue within the artificial intelligence community, encouraging researchers and developers to reflect on the role of biases in generative language models and the ongoing pursuit of ethical AI.","container-title":"First Monday","DOI":"10.5210/fm.v28i11.13346","ISSN":"1396-0466", "journal Abbreviation":"FM","note":"arXiv:2304.03738 [cs]","source":"arXiv.org","title":"Should ChatGPT

be Biased? Challenges and Risks of Bias in Large Language Models","title-short":"Should ChatGPT be Biased?","URL":"http://arxiv.org/abs/2304.03738","author":[{"family":"Ferrara","given":"Emilio"}],"accessed":{"date-parts":[["2023",12,9]]},"issued":{"date-parts":[["2023",11,7]]}}],"schema":"https://github.com/citation-style-language/schema/raw/master/csl-citation.json"}

47 Kaddour & ors. 19-20.

48 Scherer 554-561 (describing the challenges posed to AI by the 'four large V' of Big Data – Volume, Variety, Velocity, and Veracity).

49 *Ibid.* 559.

50 This is the phenomenon that judges decide more mercifully after having had lunch, see Scherer 558, 570.

51 *Ibid.* 542 (calling AI a 'black box' of legal decision-making).

52 For an overview of methods to improve the explanation of results produced by LLM, see Zhao, Haiyan and others. "Explainability for Large Language Models: A Survey" arXiv (28 November 2023) http://arxiv.org/abs/2309.01029.

53 The expression is taken from Bender, Emily M. and others. "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (ACM 2021) https://dl.acm.org/doi/10.1145/3442188.3445922.

54 See Virues Carrera, Mauricio. "Accommodating Kleros as a Decentralised Dispute Resolution Tool for Civil Justice Systems: Theoretical Model and Case of Application" (providing as annexes all documents, briefs and the award in this case); Sharma, Chandrika. "Blockchain Arbitral Award: Potential Challenges in Recognition and Enforcement under the New York Convention" (2022) 16 *Romanian Arbitration Journal/Revista Romana de Arbitraj* 85, 89-90.

55 On AI tools used in arbitration, see e.g. Berardicurti, Bianca. "Artificial Intelligence in International Arbitration: The World Is All That Is the Case" (2021) *40 under 40 International Arbitration* 377-392; Chan, Elizabeth, Nasir Gore, Kiran & Jiang, Eliza. "Harnessing Artificial Intelligence in International Arbitration Practice" (2023) 16 *Contemporary Asia Arbitration Journal* 263-300; Carrara, Cecilia. "Science and Arbitration, The Impact of Cognitive Science and Artificial Intelligence on Arbitral Proceedings – Ethical issues" (2020) *Austrian Yearbook on International Arbitration* 513, 519-527; Cabrera Colorado, Orlando Federico. "The Future of International Arbitration in the Age of Artificial Intelligence" (2023) 40 *Journal of International Arbitration* 301-342; Rajendra, Josephine Bhavani & Thuraisingam, Ambikai S. "The Deployment of Artificial Intelligence in Alternative Dispute Resolution: The AI Augmented Arbitrator" (2022) 31 *Information & Communications Technology Law* 176, 181-182. For data on the use of AI tools in arbitration see BCPL, '2023 Annual Arbitration Survey: The Rise of Machine Learning', https://www.bclplaw.com/a/web/tUW2SW6fjHrpXVrA7AfWkS/102932-arbitration-survey-2023-report_v10.pdf; Queen Mary University of London and White & Case, "2018 International Arbitration Survey: The Evolution of International Arbitration" https://arbitration.qmul.ac.uk/media/arbitration/docs/2018-International-Arbitration-Survey---The-Evolution-of-International-Arbitration-(2).PDF pp. 32-33.