Journal of Dispute Resolution

Volume 2021 | Issue 2 Article 5

2021

Can Artificial Intelligence ("AI") Replace Human Arbitrators? Technological Concerns and Legal Implications

Gizem Halis Kasap

Can Artificial Intelligence ("AI") Replace Human Arbitrators? Technological Concerns and Legal Implications

Gizem Halis Kasap [*]

I. INTRODUCTION

The global coronavirus pandemic will be the defining event of 2020. The Director-General of the World Health Organization reported that it took 67 days from the first reported case to reach the first 100,000 cases, eleven days after that to hit 200,000 and then only four days more to reach 300,000.[1] Just like the spread of a pandemic, technological advances in the legal field develop with overwhelming speed. Susskind argues that, in the next two decades or less, the legal world will change more dramatically than it has over the past two centuries.[2]

Artificial intelligence ("AI") is no longer a precursor to the future—it is already here in the mainstream. Some countries, for example, have started to implement AI-based technologies into their adjudication processes. It has been reported that Estonia is currently developing an AI judge that can adjudicate small claims disputes of less than €7,000[3] and that China already has digital courts presided over by an AI judge.[4] Together with the triggering effect of such futuristic news, AI studies

---

[*] Gizem Halis Kasap is an attorney licensed to practice in New York and Istanbul, with an S.J.D from Wake Forest Law, an LL.M from Penn State Law, and an LL.B from Istanbul University School of Law. Passionate about the interplay between arbitration and technology, she specializes in technology law and dispute resolution in the context of private international law. The author would like to thank Dr. Ayca Akkayan-Yildirim of Akkayan & Yildirim Attorney Partnership for her unwavering faith in her abilities as a lawyer and human being. The author would also like to thank the Journal of Dispute Resolution executive board and staff members for preparing this article for publication. Finally, the author would like to extend a hearty thank you to her loving and encouraging family, especially her husband, Atilla, for all her invaluable support, encouragement, and advice.

1. Bill Chappell, *Coronavirus: WHO Head Says Nations Must Attack As 'Pandemic Is Accelerating'*, NPR (Mar. 23, 2020) https://www.npr.org/sections/coronavirus-live-updates/2020/03/23/820290984/coronavirus-who-head-says-nations-must-attack-as-pandemic-is-accelerating.

2. Richard Susskind, TOMORROW'S LAWYERS, at xvii (2nd ed., 2017). Other studies supporting the argument that the legal world will dramatically change over the next two decades. *E.g., Deloitte Insight: Over 100,000 legal roles to be automated*, LEGAL IT INSIDER (Mar. 16, 2016), https://www.legaltechnology.com/latest-news/deloitte-insight-100000-legal-roles-to-be-automated/. (arguing that in the next two decades, 40% of the legal profession may be automated and replaced by artificial intelligence).

3. Eric Niiler, *Can AI Be a Fair Judge in Court? Estonia Thinks So*, WIRED MAG., (Mar. 25, 2019). https://www.wired.com/story/can-ai-be-fair-judge-court-estonia-thinks-so/.

4. Despite the futuristic headline language of the news, the term "AI Judge" does not seem to include an algorithm that can predict the outcome of the case. Rather, the AI judge is more likely to serve as assistants, relying on the knowledge and support of actual judges. *Beijing Internet Court Launches AI judge*, CHINA DAILY (June 28, 2019), http://www.chinadaily.com.cn/a/201906/28/WS5d156cada3103dbf1432ac74.html.

that predict the outcome of litigation have stirred heated debate about the possible arrival of AI judges.[5]

Despite the fact that AI may predict case outcomes and talk of AI judges, few scientific and legal studies have investigated the prediction of legal decision-making in arbitration.[6] Inspired by these debates, and to fill this gap in legal scholarship, this article poses the question of whether AI will be able to replace arbitrators, enquires into AI's ability to predict outcomes of future cases in the context of international commercial arbitration, and scouts the potential implications if AI does, in fact, replace arbitrators.

AI has been around since the 1950s, but machine learning and deep learning are its modern symbols and the reason AI is currently in its latest phase of hype.[7] Lay or legal users may find AI algorithms complex. Lawyers, however, need to grasp at least basic terminology and have an understanding of how the algorithms work if they are to analyze the legal implications for using such algorithms. For these purposes, Section II defines AI and elaborates, in broad strokes, the steps to design and deploy an AI to predict the outcome of a case.

Section III pores over two of the most notable judicial decision prediction studies involving the use of AI. It demonstrates why these AI models fall short and are thus not readily transferable to predicting the outcome of international commercial arbitration cases. This critical analysis aims to raise awareness on the side of data analysts, scientists, and other technical experts to sufficiently address the conclusions drawn on the international commercial arbitration front when building the models for predicting the outcome of an arbitration case.

Section IV points out major concerns about AI arbitration in light of state of the art AI technology. In particular, this section discusses the major risks arising from data issues, machine bias, the opacity of AI systems, and the absence of emotional intelligence in AI. In assessing the potential impact of AI arbitration, it will analyze how the current technology applies in terms of these identified challenges in order to forecast what the future might bring.

Building on the discussion in Section IV, Section V examines the legal implications of using AI in the current legal framework of international commercial arbitration. It will set out the essential values and norms accepted in international commercial arbitration and demonstrate how an AI arbitrator might come into tension with those values and what the implications of such conflicts might be. Raising concerns involving AI arbitrators, however, should not be considered an anti-

5. *E.g.*, Eugene Volokh, *Chief Justice Robots*, 68 DUKE L. J. 1135, 1142 (2019) (exploring the idea of an AI judge); Rebecca Crootof, *Cyborg Justice" And The Risk Of Technological-Legal Lock-In*, 119 COLUM. L. REV. FORUM 233, 241 (2019) (discussing the implication of AI judges); Ahmad Alozn & Abdulla Galadari, *Can Machines Replace the Human Brain? A Review of Litigation Outcome Prediction Methods for Construction Disputes*, PROCEEDINGS OF THE "NEW PERSPECTIVES IN CONSTRUCTION LAW" INTERNATIONAL CONFERENCE 75 (Mar. 2015), https://ssrn.com/abstract=2902470.

6. *E.g.*, N.B. Chaphalkar et al., *Prediction of Outcome of Construction Dispute Claims Using Multilayer Perceptron Neural Network Model*, 33 INT'L J. PROJECT MANAGEMENT 1827 (2015) (study on predicting the outcome of disputes in the context of construction arbitration). *See generally* Paul Bennett Marrow et al., *Artificial Intelligence and Arbitration: The Computer as an Arbitrator—Are We There Yet?*, 74 DISP. RESOL. J., 35 (Oct. 2020) (discussing computer arbitrators generally in the context of U.S. domestic arbitration); Maxi Scherer, *Artificial Intelligence and Legal Decision-Making: The Wide Open?*, 36 J. INT'L ARB. 539 (2019) (exploring whether and how AI will potentially replace arbitrators).

7. The true birth of AI as we know it today began with Alan Turing's publication of "Computing Machinery and Intelligence" in 1950. A.M. Turing, *Computing Machinery and Intelligence*, 59 MIND 433, 433-51 (1950).

innovative approach. Instead, this article aims to inform the construction and operation of an AI arbitrator—if and when it is technically feasible—in a way that is acceptable for arbitration users and courts.

Section VI of this article concludes by stating that arbitral proceedings heavily depend on the multifaceted components of human intelligence and thus urges both technical experts and legal actors to study the implications of AI arbitration carefully to prevent problems that can arise due to the premature deployment of AI arbitrators.

II. A BRIEF PRIMER ON ARTIFICIAL INTELLIGENCE

There are many debates and controversies about the term AI. "Artificial" simply contrasts with "natural." However, people discern "intelligence" in many different ways, which makes the term a matter of heated debate among philosophers, psychologists, scholars, and other scientists.[8] The definition of intelligence involves certain mental activities, such as learning, reasoning, understanding, grasping truths, and the like.[9] As Scherer aptly observes, the existing literature on AI discusses the term intelligence broadly, often taking human intelligence as a reference point.[10] Defining a computer's capability to achieve intelligence and drawing the lines between artificial and human intelligence are crucial if we are to answer whether an AI arbitrator can adequately rule on a dispute and protect legal rights.

The desire to build intelligent machines is as old as humanity with evidence of it among the ancient Greeks.[11] The modern AI attempts, on the other hand, started in the 1950s with logic-based, symbolic AI models.[12] The 1970s saw the emergence

8. *See generally* Shane Legg & Marcus Hutter, *A Collection of Definitions of Intelligence*, 157 FRONTIERS IN ARTIFICIAL INTELLIGENCE AND APPLICATIONS 17 (2007) (compiling definitions of intelligence made by psychologists, AI researchers, among others).

9. *Id*.

10. Scherer, *supra* note 6, at 542. For example, the Turing test is widely considered as the most useful test to determine whether a machine has a capacity like that of a human. *E.g.,* Ashok K. Goel & Jim Davies, *Artificial Intelligence*, THE CAMBRIDGE HANDBOOK OF INTELLIGENCE 468, 479 (Robert J. Sternberg & Scott Barry Kaufman eds., 2011). In the Turing test, computers and human beings are put in typed chat sessions with human interrogators. If computers can reliably fool the interrogators into believing they are human, they pass the test. Turing, *supra* note 7, at 433-35.

11. *See e.g.*, Adrienne Mayor, GODS AND ROBOTS: MYTHS, MACHINES, AND ANCIENT DREAMS OF TECHNOLOGY 1 (2018) (elucidating the ancient Greek myths and surveying the wide range of forms of artificial life in Greek mythology).

12. *See e.g.*, John McCarthy et al., *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, *August 31, 1955*, 27 AI MAGAZINE 12, 13-14 (2006), (reproducing part of the Dartmouth summer research project); Arthur Samuel, *Some Studies in Machine Learning Using the Game of Checkers*, 3 IBM J. RES. AND DEVELOPMENT 210 (1959) (being credited as one of the first working instances of a machine-learning system). The Dartmouth Summer Research Project on Artificial Intelligence is considered the birthplace of AI. McCarthy invited a group of researchers from a multitude of disciplines to this summer workshop, and this later turned out to be was one of the first serious attempts of AI as a discipline. Stan Franklin, *History, Motivations, and Core Themes*, *in* THE CAMBRIDGE HANDBOOK OF ARTIFICIAL INTELLIGENCE 15, 18 (Keith Frankish & William M. Ramsey eds., 2014). There were numerous discussions on machine intelligence and natural language processing, and they predicted that machines could soon, as early as the next generation, replace humans and reason as effectively as humans. McCarthy et al., *supra* note 12. These promises turned out to be overstated, though they are the leaders of AI research today, and the first AI winter began in early the 1970s when funding for AI diminished. Franklin, *supra* note 12, at 21. The first AI winter ended in the 1970s when expert systems first appeared. *Id.*

of expert systems,[13] one of its subsets being rule-based systems, which is the focus of this article. Rule-based systems are hinged on a set of rules and use "if/then" statements to establish rule-of-thumb decisions.[14] A rule-based system is built on two main components: 1) a set of facts about a situation, and 2) a set of rules applicable to those facts. This requires the involvement of human experts who could specify all the steps that the machine needs to take to make a decision.[15]

These systems have great advantages. For one, they render a decision based on expert human knowledge in a readable form. As such, they provide transparency to the user and the programmer, because the human can review the work of the system and find the rules that have been used to render a decision.[16]

Though rule-based systems are reliable and transparent, they are not feasible or economically viable for certain tasks. Knowledge is learned by human experts and provided to the system. It requires consistent effort by humans to carefully gather, distill, and combine expert knowledge from the outside.[17] Keeping the knowledge base up to date by adding new knowledge or changing the rules is extremely laborious.[18] As these systems use "if/then" principles, or logical reasoning in general, they are not well-versed enough to handle complex and dynamic realities.[19] For these reasons, the 1990s saw rule-based systems lose popularity, and the next wave in AI emerged.[20] Here, machine learning was emphasized, moving from an approach focused on knowledge to one driven by data.

Machine learning refers to a science of automatic pattern recognition between variables in a dataset, often in order to predict or estimate some result.[21] A machine learning-based program defines the rules on its own through trial and error, and judges performance according to the objectives defined by the programmer.[22] At its heart, it revolves around the inference of hidden factors, patterns, correlations, and rules by observing relevant data.[23]

Machine learning requires a goal (e.g., identifying a dog in a picture) and the collection of data (e.g., dog pictures) from which to learn. From there, it invents

13. *E.g.*, Michael Kifer & Yanhong Annie Liu, DECLARATIVE LOGIC PROGRAMMING: THEORY, SYSTEMS, AND APPLICATIONS 14 (2018); ETHEM ALPAYDIN, MACHINE LEARNING 50 (2016).

14. *Id.* at 49-50; Margaret A. Boden, *Gofai*, *in* THE CAMBRIDGE HANDBOOK OF ARTIFICIAL INTELLIGENCE, *supra* note 12, at 89-91.

15. ALPAYDIN, *supra* note 13, at 50; Franklin, *supra* note 12, at 20.

16. Richard E. Susskind, *Expert Systems in Law: A Jurisprudential Approach to Artificial Intelligence and Legal* Reasoning, 49 MODERN L. REV., 168, 173 (1986).

17. ALPAYDIN, *supra* note 13, at 50; Franklin, *supra* note 12, at 20.

18. ALPAYDIN, *supra* note 13, at 51; PAMELA MCCORDUCK, MACHINES WHO THINK 425, 425 (2nd ed. 2004).

19. ALPAYDIN, *supra* note 13, at 51.

20. MCCORDUCK, *supra* note 18, at 442.

21. *E.g.*, ALPAYDIN, *supra* note 13, at 41; David Lehr & Paul Ohm, *Playing with the Data: What Legal Scholars Should Learn About Machine Learning*, 51 U.C. DAVIS L. REV. 653, 671 (2017) ("[M]achine learning refers to an automated process of discovering correlations (sometimes alternatively referred to as relationships or patterns) between variables in a dataset, often to make predictions or estimates of some outcome."). Though this article mainly focuses on machine prediction as a foundation of AI arbitrator, machine learning has many other vital uses. David Danks, *Learning*, *in* THE CAMBRIDGE HANDBOOK OF ARTIFICIAL INTELLIGENCE, *supra* note 12, at 151, 157 ("As with human learning, the value of machine learning is less in the output, and more in the way that the output can be used for future tasks: prediction, planning, classification, recognition, and so on.").

22. Lehr & Ohm, *supra* note 21, at 671-72.

23. Harry Surden, *Machine Learning and Law*, 89 WASH. L. REV. 87, 91 (2014) ("The goal . . . is to build an internal computer model of some complex phenomenon . . . that will ultimately allow the computer to make automated, accurate classification decisions.").

rules that help to accomplish its goal using trial and error in any given situation. Different than rule-based programs, the programmer does not hand-code the algorithms into the machine. After looking at data, the machine detects and extracts the necessary algorithms to achieve the goal (e.g., identifying a dog in the picture correctly) defined by the programmer. As such, the programmer does not need to write a code that defines what a dog is. In this sense, machine-learning algorithms are said to be like a baby that learns through examples, whereas rule-based algorithms are similar to a human being born with fixed knowledge.[24]

It is no surprise, then, that machine learning can identify rules and connections that were not even known to the programmer. For this reason, it is an attractive problem-solver and particularly useful, especially where the rules are genuinely complicated or mystifying. Through machine learning, people can predict the future to a certain extent, categorize objects in a meaningful way, carry out complex analysis, or complete mundane tasks by automation.

The fundamental methods used to design and deploy machine-learning technology to predict the outcome of a case can be broken down into steps.[25] The first step is to collect data and transform it into a format that can be used by machine-learning technology. The AI's hunger for data is the main reason that the "big data" era—where people collect and analyze large sets of data—has gone hand in hand with the AI era.[26] The ability to collect vast volumes of data on the web and the efficiency of computing to process that data have made the creation of data solutions

---

24. Janelle Shane, an artificial intelligence researcher, illustrates the difference between the rule-based programming and machine-learning through knock-knock jokes. JANELLE SHANE, YOU LOOK LIKE A THING AND I LOVE YOU 9 (2019). First, she analyzes the formula of the knock-knock rules, which is as follows: "Knock, knock. Who's there? [Name]. [Name] who? [Punchline]." *Id.*
She then provides a list of valid names and a list of valid punchlines to computer to produce knock-knock jokes. *Id.* at 10. This way the computer can create knock-knock jokes by selecting and slotting the name-punchline-pair from the list provided by programmer Shane (*e.g.,* if the name is "*lettuce*", then the punchline is "*in, it's cold out here!*"). *Id.* The computer doesn't create new knock-knock jokes, it generates logical statements based on user-introduced premises and pre-programmed logical connections and conclusions. *Id.* Shane then moves forward to illustrate machine learning through the knock-knock jokes. *Id.* at 11. She gives a set of existing knock-knock jokes and instructs the compute to generate more knock-knock jokes using random letters and punctuation marks. *Id.* The computer uses a trial-and-error method to generate passable knock-knock jokes. *Id.* at 12-13 First, it learns about the rules about which order the letters come in (such as guessing that "*o"* is often followed by "*ck"*) and then figures out the rest of the formula and finally produces an actual joke that is not on the set of existing jokes provided by Shane, which is as follows: "Knock Knock. Who's There? Alec. Alec who? Alec-Knock Knock jokes." *Id.* at 17.

25. Charles Ciumei QC, *Digital Justice and The Use of Algorithehms to Predict Litigation Outcomes*, EUROPEAN CIRCUIT OF THE BAR ANNUAL CONFERENCE 2 (Sep. 21, 2018) (revised version available in Essex Court Chambers website); Lehr & Ohm, *supra* note 21, at 655 ("Far from a straight linear path, most machine learning dances back and forth across these steps, whirling through successive passes of model building and refinement.").

26. ALPAYDIN, *supra* note 13, at 166. "Data is the new oil" is a well-known saying that is often attributed to Clive Humby. Michael Palmer, *Data is the New Oil*, ANA MKTG. MAESTROS (Nov. 3, 2006, 5:43 AM), http://ana.blogs.com/maestros/2006/11/data_is_the_new.html. The metaphor reflects two apparent parallels. The first and most obvious of which is that data, like oil, is an *extremely valuable commodity*. *Id.* It also suggests that data, like oil, requires *significant processing* if any of the value is to be realized, and this processing is called as data analysis. *Id.* Data analysis consists of four main dimensions; (i) cleaning, (ii) transforming (iii) modeling, and (iv) inspecting. Prateek Bihani & S. T. Patil, *A Comparative Study of Data Analysis Techniques*, 2 INT'L J. EMERGING TRENDS & TECH. IN COMP. SCI. 95, 95 (2014).

possible.[27] For these reasons, AI has come to play a big part in our daily lives in the past two decades.[28] Analyzing this data generates a quality input for machine learning algorithms to use and generate correct results.[29]

The second step is to choose an applicable model. Algorithms excel at addressing specific problems, and thus one algorithm can be proven superior to another only in relation to a specific problem. The programmer therefore needs to choose an appropriate machine-learning model that can yield the best result for the specific machine learning problem; for example, predicting the outcome of an arbitration case.[30] Many open-source machine learning models are available off-the-shelf to researchers and businesses, as well as proprietary ones.

After choosing an applicable model, the next step is to train the AI model. In the current state of the technology, solving a small problem with AI requires extensive training with data sources.[31] Despite their common need for large volumes of data, algorithms can learn from data in very different ways. For the time being, unsupervised learning algorithms have the least legal value. Supervised learning algorithms are, therefore, receiving the most legal attention.[32]

In supervised learning, the algorithm is given an input and an output, and the goal is to develop a map from the input to the output where the supervisor provides

---

27. ALPAYDIN, *supra* note 13, at 10-11; Peter Stone et al., *Artificial Intelligence and Life in 2030: One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel* 51 (2016), https://ai100.stanford.edu/2016-report. [hereinafter *Artificial Intelligence and Life in 2030*].

28. ALPAYDIN, *supra* note 13, at ix; Stone et al., *supra* note 27, at 51.

29. Lehr & Ohm, *supra* note 21, at 669. So far, scientific studies to predict the outcome of judicial cases have used the judgments in prior cases as raw material. *See generally* Z. Liu & H. Chen, *A predictive performance comparison of machine-learning models for judicial cases*, 2017 IEEE SYMP. SERIES ON COMPUTATIONAL INTELLIGENCE 1 (2017). These prior cases have been classified and/or portioned based on the parts that the researchers identified as having predictive value (*i.e.,* features). The set of features may include non-textual information *(e.g*., who the judge is, procedural history, and term) or the semantic of law and case texts. *See e.g*., Daniel Martin Katz et al., *A general approach for predicting the behavior of the Supreme Court of the United States*, 12 PLOS ONE 1, 5 (2017), https://doi.org/10.1371/journal.pone.0174698 (identifying features as justice, term, natural court, month of argument, and the like). *Id.* Nikolaos Aletras et al., *Predicting Judicial Decisions of the European Court of Human Rights: A Natural Language Processing Perspective*, 2 PEERJ. COMPUT. SCI. 92 (2006), https://peerj.com/articles/cs-93/ (using the textual content extracted from a case as input).

30. There is no single algorithm that best fits all conditions and data sets imaginable, which is known as the no free lunch theorem. John D. Kelly her, et al., *Fundamentals of Machine Learning for Predictive Data Analytics 13* (1st ed. 2015). In fact, there are five main camps of machine learning: Symbolists, Connectionists, Evolutionaries, Bayesians, and Analogizers. Pedro Domingos, *The Master Algorithm: How The Quest for The Ultimate Learning Machine Will Remake Our World 51* (2015). Each of these tribes has a different technique and strategy for solving machine problems. *Id.* at 53. For purposes of predicting the outcome of future cases, scientists have used a variety of techniques with success. *E.g.*, Katz et al., *supra* note 29, at 6 (using random forest from the Symbolists tribe); Aletras et al., *supra* note 29, at 2 (using Support Vector Machines from the Analogizers tribe). Nonetheless, machine learning's end goal is to integrate the techniques of the five tribes so that they can develop a single algorithm (a master algorithm) that can solve all AI problems. Domingos, *supra* note 30, at 51.

31. ALPAYDIN, *supra* note 13, at 24-25.

32. Lehr & Ohm, *supra* note 21, at 673, 676. Unsupervised learning, on the other hand, uses data that is not labeled. Hence, the algorithm decides and discovers the data patterns on its own. ALPAYDIN, *supra* note 13, at 111; Danks, *supra* note 21, at 154. As in unsupervised learning, reinforcement learning uses examples that are not labeled. What is different, however, the trainer gives positive or negative feedback depending on the solution that the algorithm presents. Overall, training ends after the algorithm reaches a satisfactory level of accuracy. ALPAYDIN, *supra* note 13, at 125-28; Eduardo Alonso, *Actions and agents*, *in* THE CAMBRIDGE HANDBOOK OF ARTIFICIAL INTELLIGENCE, *supra* note 12, at 232, 236.

correct values during the training to initiate learning.[33] By learning from data that is already labeled with desired answers, the algorithm can more readily solve a given problem, like predict the outcomes of unforeseen arbitration cases. As such, the value provided by the supervisor enables the algorithm's output to get as close as possible to these desired outputs.[34] Finally, once the algorithm is trained, it can be used to predict decisions that the algorithm has not seen, which in this article is undecided arbitral cases.

III. THE PROMISE OF AI ARBITRATORS: TAKING STOCK OF THE SCIENCE

The study of court-decision predicting has a longer history in the United States than in other parts of the world.[35] The overwhelming number of these studies use the U.S. Supreme Court's database to predict future decisions by the Court and the behavior of Justices considering their political background.[36] In countries other than the United States, studies on court-decision predicting are not very common, although a few initiatives exist that analyze cases before international courts.[37] Reviewing scientific studies on the prediction of legal decision-making is of importance, as these may lend support to the development of AI arbitrators and shed light on how an AI arbitrator might operate. The next two subsections review the most notable studies in this area and their implications for arbitration.

A. Predicting the Decisions of the U.S. Supreme Court

An oft-cited 2017 study by Katz and others on the prediction of legal decision-making draws on data from the U.S. Supreme Court decisions. The study aims to create a generalized methodology applicable to all U.S. Supreme Court decisions over time without focusing on an individual Supreme Court Justice or a given term.[38] In this study, all of the Supreme Court's decisions between 1816 and 2015 were analyzed. The data consisted of 28,009 cases and 243,882 votes by individual judges.[39] A notable aspect of this study, therefore, is that, unlike earlier studies with

33. ALPAYDIN, *supra* note 13, at 38-40; Danks, *supra* note 21, at 154; Lehr & Ohm, *supra* note 21, at 673-677.

34. ALPAYDIN, *supra* note 13, at 38.

35. Medvedeva et al., *Judicial Decisions of the European Court of Human Rights: Looking into the Crystal Ball*, PROCEEDINGS OF THE CONFERENCE ON EMPIRICAL LEGAL STUDIES IN EUROPE 2018, at 1, 3.

36. *E.g.*, Lee Epstein, The Behavior of Federal Judges: A Theoretical and Empirical Study of Rotional Choice (Harvard Univ. Press 2013); N. Sivaranjani et al., *A Broad View of Automation In Legal Prediction Technology*, PROCEEDINGS OF THE THIRD INTERNATIONAL CONFERENCE ON ELECTRONICS COMMUNICATION AND AEROSPACE TECHNOLOGY, 2019, at 180, 184 [hereinafter ICECA 2019] (surveying the existing studies on case outcome prediction and finding that 80% of the models are developed to predict the behavior of the U.S. Supreme Court); Jeffrey J. Rachlinski & Andrew J. Wistrich, *Judging the judiciary by the numbers: Empirical research on judges*, 13 ANN. REV. L. SOC. SCI. 2017, at 203; Theodore W. Ruger, et al., *The Supreme Court Forecasting Project: Legal and Political Science Approaches to Predicting Supreme Court Decisionmaking*, 104 COLUM. L. REV. 1150 (2004).

37. Sivaranjani et al., *supra* note 36, at 183; Medvedeva et al., *supra* note 35, at 3.

38. Katz et al., *supra* note 29, at 1. The successor studies were also considered to be successful, but they have limited transposability to a general U.S. Supreme Court prediction. *See e.g.,* Ruger, et al., *supra* note 36, at 1150 (limiting its data with the cases of the Rehnquist U.S. Supreme Court since 1994 and before the 2002 Court term).

39. Katz et al., *supra* note 29, at 5.

quite limited data, a large volume of training data and complex machine learning technology were applied.[40]

The researchers prepared a model by selecting the different set of input features, including the features related to the background and substance of the cases;[41] the Circuit Court of Appeals from which the dispute arose;[42] chronologically oriented aspects of the oral argument and case timing;[43] and the behavior of Justices.[44]

After learning from the sample drawn from the dataset, the algorithm was applied to the remaining, out-of-sample data, and asked to make predictions regarding two questions: 1) whether the Court as a whole would affirm or reverse a judgment, and 2) how each Justice would vote.[45] Demonstrating a high rate of accuracy compared to its predecessors, the model correctly predicted 70.2 percent of the U.S. Supreme Court's decisions and 71.9 percent of the Justices' votes.[46]

Despite its success, several questions regarding the U.S. Supreme Court study remain unaddressed, and the applicability of this study to arbitration appears limited. First, as Scherer aptly put it in his recent article on the topic, it is questionable whether the model can be readily adapted to lower court decisions, given such courts function to decide a dispute rather than to act as an appellate body.[47] This is because the majority of the features in training data are related to background and chronological aspects of the case as well as the Justices' behavior, whereas there are a limited number of features related to the merits of the dispute.[48]

Moreover, the study limited itself to cases where the U.S. Supreme Court reviewed lower court decisions. Thus, it does not include in its training data cases where the Court has original jurisdiction.[49] This is because "the Court's decision might otherwise result in a complex outcome that does not map onto a binary outcome" as to whether the lower court's decision is reversed or affirmed.[50] Like lower

40. Sivaranjani et al., *supra* note 36, at 182 ("[K]atz model is considered to be efficient because the previous research considers only the short span of cases.").

41. Katz et al., *supra* note 29, at 5 (including petitioner, respondent, manner in which the court takes jurisdiction, administrative action, issue, and issue area, among other features).

42. *Id*. (including the opinion under review, the origin corresponds to the location of original filing and the like).

43. *Id*. ("These features include (i) whether or not oral arguments were heard for the case, (ii) whether or not there was a rehearing, and (iii) the duration between when the case was originally argued and a decision was rendered.").

44. *Id*. at 6. ("These features fall into three categories: (i) features related to the rate of reversal, (ii) features related to the left-right direction of a decision, and (iii) features related to the rate of dissent.").

45. *Id*. at 2.

46. Sivaranjani et al., *supra* note 36, at 182; Katz et al., *supra* note 29, at 8.

47. Scherer, *supra* note 6, at 552-53

48. *Id*.; Katz et al., *supra* note 29, at 5.

49. Katz et al., *supra* note 29, at 4.

50. Scherer, *supra* note 6, at 553; Katz et al., *supra* note 29, at 4. *See also* André Lage-Freitas et al., *Predicting Brazilian court decisions* 1, 2 (Apr. 20, 2019), https://arxiv.org/abs/1905.10348. On the study predicting the Brazilian Supreme Court's decisions on appeal, the researchers identified six different classifications for the decisions as (i) *not-cognized* when the court did not accept to hear the appeal request; (ii) *yes* for full favorable decisions; (iii) *partial* for partially favorable decisions; (iv) *no* when the appeal was denied; (v) *prejudicada* to mean that the court could not hear the case due to an impediment such as the appellant died and the like; and (v) *administrative* when the court acts as courts of conflict to resolve the conflict of competence between lower court judges. *Id.* Similar to the U.S. Supreme Court, the researchers then removed the decisions that are classified as *prejudicada*, *not-cognized* and *administrative,* stating that "these labels refer to very peculiar situations which are not useful for prediction purposes." *Id.* Thus, similar to the U.S. Supreme Court, they did not include rare instances. This raises the question of whether AI models can replace the judges or arbitrators entirely, as they are

courts, arbitral tribunals are tasked with resolving a dispute, as opposed to reviewing the decision of another court or tribunal. Moreover, the disputes in international arbitration tend to involve complicated and intricate matters of fact and law that can be difficult to map into a binary outcome model. Hence, it remains an open question whether an AI model can predict a case outcome when the facts are highly complex and not easily converted to a binary classification.[51]

Second, instead of focusing entirely on the internal content of the cases and the actual court materials, the model learned from features related to the political leanings of the Justices, among other factors.[52] While across the United States, judicial appointments are highly partisan and partisan leanings often guide judges' decisions, the usual method of judicial selection in the rest of the world is not necessarily political.[53] In this respect, the study's goal to build a general model appears not readily transferable to other jurisdictions.

Turning to arbitration, one can argue that arbitrators' political leanings come into play more in the context of investment arbitration. International commercial arbitration, instead, is more fact-driven.[54] Thus, certain features related to judges' political orientation—which are central to the cited study model—are less likely to be determinative in international arbitration and thus might not transfer readily to that domain.[55]

B. Predicting the Decisions of the European Court of Human Rights

In another study that has received widespread acclaim, a group of researchers focused on the decisions of the European Court of Human Rights ("ECtHR"). The researchers trained machine learning algorithms using previously decided cases on certain articles of the European Convention for Human Rights: Article 3 prohibiting torture, Article 6 protecting the right to a fair trial, and Article 8 protecting the right

---

usually designed for a specific purpose and are not models intended for a broad application. Marrow et al., *supra* note 6, at 52.

51. Different than the U.S. Supreme Court study, the study on predicting the Brazilian Supreme Court decisions does not conduct a binary task and instead uses three possible prediction results consisting of *yes*, *partial,* and *no*. Lage-Freitas et al., *supra* note 50, at 2. In this sense, it appears to be more promising in handling complex cases. Arguably, the next step could be making a probability distribution so that one can forecast with what likelihood a party would win the cases given a feature. *See* Daniel L. Chen & Jess Eagel, *Can Machine Learning Help Predict the Outcome of Asylum Adjudications?*, PROCEEDINGS OF THE ACM CONFERENCE ON AI AND THE LAW 1, 9 (2017).

52. Katz et al., *supra* note 29, at 6.

53. *E.g.*, JEFFREY A. SEGAL & HAROLD J. SPAETH, THE SUPREME COURT AND THE ATTITUDINAL MODEL, 65 (1993); DAVID ROHDE & HAROLD SPAETH, SUPREME COURT DECISIONS MAKING 72 (1976); Neal Devins & Lawrence Baum, *Split Definitive: How Party Polarization Turned the Supreme Court into A Partisan Court*, 2016 SUP. CT. REV. 301, 302 (2016); Adam Liptak, *U.S. voting for judges perplexes other nations*, NY TIMES, (May 25, 2008) https://www.nytimes.com/2008/05/25/world/americas/25iht-judge.4.13194819.html. *See also* Cass Sunstein et al., *Ideological Voting on Federal Courts of Appeals: A Preliminary Investigation*, 90 VA L REV 301 (2004) (finding a correlation between judges' attitude and the ideology of the official who appointed a judge in various legal areas).

54. *E.g.*, Jason Webb Yackee, *Controlling the International Investment Law Agency*, 53 HARV. INT'L L. J. 391, 393-94 (2012) (noting that cases require "complex and politically fraught value-balancing exercises"); Michael Waibel & Yanhui Wu, *Are Arbitrators Political?* 40 (Dec. 13, 2011), available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2101186 (unpublished manuscript) (cited with permission) (conducting an empirical study and finding that investment arbitrators appear to be influenced by their policy views and do not merely apply the law as it stands when deciding on investment disputes).

55. Scherer, *supra* note 6, at 553-54.

to respect for private and family life.[56] The study used an equal number of decisions finding a violation of the Convention and those finding no violation for each article.[57] Like the U.S. Supreme Court study, the goal output was based on a binary classification—was there a violation of the respective articles of the Convention or not?[58]

What is most notable in this study is that, unlike the U.S. Supreme Court study, the model was based exclusively on the textual content of the actual decisions without regard to other background or chronological variables or the Justices' political leanings. The researchers indicated that the data used in the study was the text extracted from the procedure,[59] facts,[60] and law[61] subsections of the decisions, excluding the operative provisions,[62] where the Court reveals the outcome of the case.[63] After the training, the model predicted the correct outcome with 79 percent accuracy using the whole text.[64] The researchers also noted that an accuracy of 73 percent was reached using only the part where the factual circumstances are described.[65]

Despite its high success rate, this study suffers from serious weaknesses. First, the researchers did not have access to other case documents apart from the judgments themselves. As such, the predictive tasks were only based on the text of the published judgments rather than the applications lodged with the Court or the briefs.[66] The portion of the decision that contains a judges' legal reasoning is not available to the parties before the trial and thus is not helpful in predicting a judgment in advance.[67] Therefore, the model turned one end product into another end product, one judgment into another judgment. Rather than forming a new end product from the raw components, the applications, briefs, and other salient parts of a judgment. This raises important questions in terms of the applicability of the model for ex-ante decision-outcome prediction.[68]

Second, the judgments used in the study were already crafted in a way to justify the finding. As such, the texts of the judgment are considered biased data and are thus problematic for ex-ante outcome prediction.[69] Moreover, the phrase "legal

56. Aletras et al., *supra* note 29, at 8.

57. *Id*.

58. *Id*. at 9.

59. *Id*. at 4. ("This section contains the procedure followed before the Court, from the lodging of the individual application until the judgment was handed down.").

60. *Id*. ("This section comprises all material which is not considered as belonging to points of law, i.e., legal arguments.).

61. *Id*. at 5. ("The law section considers the merits of the case, through the use of legal argument.").

62. *Id*. at 6. ("This is the section where the Court announces the outcome of the case, which is a decision to the effect that a violation of some Convention article either did or did not take place.").

63. *Id*. at 10-11.

64. *Id*. at 1.

65. *Id*. at 10.

66. Aletras et al., *supra* note 29, at 5 (The researchers' assumption was "that the text extracted from published judgments of the Court bears a sufficient number of similarities with, and can, therefore, stand as a (crude) proxy for, applications lodged with the Court, as well as for briefs submitted by parties in pending cases.").

67. Scherer, *supra* note 6, at 549-50. Medvedeva et al., *supra* note 35, at 10.

68. Scherer, *supra* note 6, at 549-50; Frank Pasquale & Glyn Cashwell, *Prediction, Persuasion, And The Jurisprudence of Behaviourism*, 68 UNIV. OF TORONTO L. J. 63, 70 (2018) (noting that a truly predictive algorithm would use the data that was in existence before the judgment itself).

69. *E.g.*, Scherer, *supra* note 6, at 550; Pasquale & Cashwell, *supra* note 68, at 70; Mireille Hildebrandt, *Algorithmic regulation and the rule of law*, PHIL. TRANS. R. SOC. A, May 2018, at 1, 7.

reasoning" speaks for itself. Data drawn from the law section contains the judges' legal reasoning in reaching the finding of a violation or no violation, and even on occasion explicit language stating the decision. For example, consider the following text extracted from the case of *Velcheva v. Bulgaria*, one of the cases cited in the study,[70] "Accordingly, there has been a violation of Article 6 § 1 of the Convention."[71]

It is unclear from the study whether researchers explicitly excluded such language from the analysis, but it appears that they did not.[72] Any layperson, much less a lawyer, reading this quote will be able to answer whether there was a violation of the right to a fair trial under Article 6 of the European Convention on Human Rights.[73] The texts of the judgment containing arguments and discussions will eventually suggest whether there is a violation or not. Therefore, using these inherently biased data is questionable for ex-ante outcome prediction.[74]

Finally, the model drew its conclusions based on the Court's articulation of the facts of the case, rather than the parties' own characterizations of the facts, which is also problematic for ex-ante outcome prediction. Indeed, the researchers accept this limitation by stating that the ECtHR's fact compositions might be customized to suit a desired outcome but nonetheless conclude that the facts of the case have the highest predictive value.[75] This view is also supported by Morison and Harkens, who argued the facts are not in dispute at this stage given the ECtHR is as an appeals court, and thus raise no objection to this.[76] Pasquale and Cashwell, meanwhile, wryly suggest that what the study did was akin to "predicting" what a food is by looking at its nutrient composition.[77]

While it can be argued that the ECtHR is bound by the conclusions of fact established in the domestic courts,[78] the ECtHR judges may nonetheless highlight

70. Aletras et al., *supra* note 29, at 6.

71. Velcheva v. Bulgaria, App. No. 35355/08, [2015] ECHR 552, para. 48 (2015).

72. *Compare* Aletras et al., *supra* note 29, at 5 (including texts from the Law subsection which includes the Court's legal reasoning) and *Id*. at 8 (stating that ". . . any sections on operative provisions of the Court are excluded. In this way, we ensure that the models do not use information pertaining to the outcome of the case," which support the argument that only the operative provisions are excluded from the training data). *See also* Scherer, *supra* note 6, at 549-50 (concluding that only the operative sections of the decisions are excluded).

73. Even if we assume for a moment that Aletras and others excluded such explicit languages from the Law part, there are still languages implying a high possibility of violation. Take the following quote from the same decision: "The Court sees no reason not to conclude that the refusal was also contrary to the requirements of the Convention set out above, since it was not permissible for the Department, a State body, to call into question the findings of a final court judgment." Velcheva v. Bulgaria, App. No. 35355/08, at para. 40. A lawyer, or even a layperson, would have no problem discerning that this language created a very high possibility for violation of the right to a fair trial.

74. *E.g.*, Scherer, *supra* note 6, at 549. In another study, Medvedeva and others also used the textual content of the ECtHR decisions but refrained from including the Law part of the judgment. The researches were aware of this drawback and thus explicitly noted that "[w]e . . . removed the Law part of the judgement as it includes arguments and discussions of the judges that partly contain the final decisions. . . . if we let our programme predict the decision based on this information, it will be unfair as the text already shows the decision." Medvedeva et al., *supra* note 35, at 10-11.

75. Aletras et al., *supra* note 29, at 5.

76. John Morison & Adam Harkens, *Re-engineering justice? Robot judges, computerized courts and (semi) automated legal decision-making*, 39 LEGAL STUD. 618, 632 (2019).

77. Pasquale & Cashwell, *supra* note 68, at 69.

78. Garcia Ruiz v. Spain, 1999-I EUR. CT. H.R. 87 (1999) (applying the so-called fourth-instance doctrine and finding that the ECtHR's function is not to deal with errors of fact allegedly committed by a national court unless and in so far as they may have infringed rights and freedoms protected by the Convention); PAUL W. KAHN, MAKING THE CASE: THE ART OF THE JUDICIAL OPINION 137 (2016) (". . .

those facts that are relevant to justifying their findings (cherry-picking), without including the other non-relevant facts that the parties pleaded (lemon-dropping).[79] Thus, the high accuracy rates do not reflect correct predictions, but instead, identify a mere correlation or relationship between the facts of the case and the outcome.

Despite its shortcomings, the implications of the ECtHR study are significant concerning AI arbitration. This is because, unlike the U.S. Supreme Court study, the ECtHR research drew on the actual text of judgments using natural language processing without background or chronological variables or information about the Justices' behavior. Thus, it seems to apply to the prediction of arbitral awards.[80] It has been noted that to predict the outcome of arbitration cases accurately the necessary data set would ideally include transcripts from actual arbitration proceedings, actual awards, all known reported judicial opinions issued by courts embodying the complete state of arbitration jurisprudence, all relevant statutes and rules of the arbitration process, and all salient journal and law review materials.[81] In reality, in predicting the outcome of decisions, a model draws only on previous cases as data, at least for the time being.[82] Therefore, until and unless a "comprehensive" model is developed, and a database for it created, actual awards and their internal content remain important for prediction purposes.

On the other side of the coin, the researchers noted that they were able to use the textual information extracted from the ECtHR judgments because the sections covering the contents were clearly separated, allowing for easy standardization and thus making text-based analysis possible.[83] Although international arbitral awards, as opposed to domestic awards, are more detailed, what forms a reasoned award in the international commercial arbitration context or how to write such an award is less clear.[84] Assuming that a model would be based on the textual information contained in arbitral awards, among other factors, distinguishing between factual findings and legal conclusions is likely to pose challenges.[85]

---

an appellate court may be even more constrained than a trial court, for it is bound by the conclusions of fact established in the record of the lower court.").

79. *E.g.*, Scherer, *supra* note 6, at 539; Simon Stern, *Introduction: Artificial Intelligence, Technology, and The Law*, 68 UNIV. OF TORONTO L.J. 1, 6-7 (2018) (". . . the study draws on the descriptions offered by the court, in the process of explaining its decision — a detail that is endogenous to the very outcome being predicted."); Pasquale & Cashwell, *supra* note 68, at 70.

80. *See* Young-Yik Rhim & KyungBae Park, *The Applicability of Artificial Intelligence in International Law*, 12 J. EAST ASIA & INT'L L. 7, 21-22 (2019) (noting that the study's significance, as it did not consider the subjective characteristics of the judges).

81. Marrow et al., *supra* note 6, at 36.

82. *See e.g.*, Lage-Freitas et al., *supra* note 50 (using the decisions of the Brazilian Supreme Court); Katz et al., *supra* note 29 (using the cases decided by the U.S. Supreme Court between 1816-2015); Aletras et al., *supra* note 29, at 3 (solely using cases retrieved through the electronic database of the ECtHR); Chaphalkar et al., *supra* note 6, at 1828 (using arbitral awards); *See generally* Scherer, *supra* note 6, at 539 (discussing the AI prediction of arbitration outcome with the assumption that arbitral awards will be the input data); *See also* Rhim & Park, *supra* note 80, at 30 (noting that securing the related materials in international arbitration is difficult).

83. Aletras et al., *supra* note 29, at 4.

84. S. I. Strong, *Reasoned Awards in International Commercial Arbitration: Embracing and Exceeding the Common Law-Civil Law Dichotomy*, 37 MICH. J. INT'L L. 1 (2015).

85. *See id*. at 37 (discussing the ambiguities in forming arbitral awards). This obstacle has not been solved in terms of litigation, either. The court decisions form an unstructured data in terms of extracting relevant information because of the Judges' differing approaches in opinion writing. Benjamin Alarie et al., *How artificial intelligence will affect the practice of law,* 68 TORONTO L. J. 106, 118 (2018). *See also* Miller v. Fenton, 474 U.S. 104, 113–14 (1985) ("Perhaps much of the difficulty . . . stems from the

Another related problem stems from differences in the approach taken by civil law and common law courts in writing awards. The former reasons from principles to instances, the latter from instances to principles.[86] Therefore, even if awards are clearly divided into different thematic sections, nature of the content may nonetheless differ depending on whether the arbitrator is trained in a civil or common law jurisdiction.

Finally, just as in the U.S. Supreme Court study, the data in the ECtHR study came from an apex court that acts as a court of last resort. International commercial arbitration, however, is more fact-driven, resolving both factual and legal disputes, as opposed to reviewing the decision of another court or tribunal. Therefore, we again face the question of whether an AI model can predict the outcome of a case when the case is not easily converted to a binary classification such as violation/no violation, affirm/reverse, and so on.

Overall, despite the general success of the U.S. Supreme Court and ECtHR studies, it appears that they are not readily applicable to arbitration, and the model will likely continue to fall short in predicting case outcomes in this domain. Additional studies are required to understand more completely the critical tenets of outcome prediction in international arbitration.

IV. AI ARBITRATION IN LIGHT OF THE TECHNOLOGICAL STATE OF THE ART

With the advent of "big data" in recent years, the importance of data in AI has grown. Lehr and Ohm argued that current scholarship is overly focused on the implications of the running model, thereby overlooking most of the advantages and disadvantages that result from the data itself.[87] Indeed, the vast majority of scholars emphasize the significance of data, stating that AI algorithms are only as good as the data they draw on.[88] This section, therefore, looks at the limitations of AI arbitration as far as data is concerned.

A. Lack of Data

As with all AI systems, the efficiency of an AI arbitrator will heavily depend on the quantity of data that is fed into it.[89] An AI arbitrator will need to analyze

practical truth that the decision to label an issue a "question of law," a "question of fact," or a "mixed question of law and fact" is sometimes as much a matter of allocation as it is of analysis.").

86. Thomas Mackay Cooper, *The Common and the Civil Law – A Scot's View*, 63 HARV. L. REV. 468, 471 (1950); *See also* Strong, *supra* note 84, at 37 ("arbitrators from common law jurisdictions often spend a significant amount of time discussing the underlying facts and analyzing legal precedents while arbitrators from civil law jurisdictions focus more heavily on categorizing the type of legal issues at stake during the initial stages of the analysis.").

87. Lehr & Ohm, *supra* note 21, at 655.

88. *E.g.*, Marrow et al., *supra* note 6, at 49; Ignacio N. Cofone, *Algorithmic Discrimination is an Information Problem*, 70 HASTINGS L.J. 1389, 1402 (2019); Scherer, *supra* note 6, at 559; Lehr & Ohm, *supra* note 21, at 712; Surden, *supra* note 23, at 106.

89. *E.g.*, Karl Manheim & Lyric Kaplan, *Artificial Intelligence: Risks to Privacy and Democracy*, 21 YALE J. L. & TECH. 106, 122 (2019) (noting that the larger the accessible data set is, the better capable AI will be); Marrow et al., *supra* note 6, at 43; Scherer, *supra* note 6, at 554; *See also* Alarie et al., *supra* note 85, at 120 (noting that predicting how courts would decide is inappropriate where data is insufficient); ARTIFICIAL INTELLIGENCE AND LIFE IN 2030, *supra* note 27, at 22 (implying that the availability of large-scale data makes a domain an ideal candidate for machine learning applications). Researchers

enough historical data to form a general rule that can be applied to novel scenarios. The volume of training cases will play a significant role in the accuracy of any prediction.[90]

While the term "big data" connotes vast data sets, finding data in such quantities is doubtful in the context of international arbitration.[91] First and foremost, international arbitral awards are seldom published and those that are published are mostly available in a heavily redacted form.[92] That is because of the nature of arbitration itself. Its privacy and confidentiality are widely seen as one of the critical advantages of arbitration over litigation.[93] Yet, it is also the source of enduring questions as to the legitimacy of arbitration, especially in the context of investment arbitration.[94] This criticism has also affected international arbitration practice, so much so that a growing chorus has called for arbitration to be made more transparent and for awards to be published.[95] Nonetheless, when compared with litigation, access to arbitral awards and materials is severely limited. Thus, compiling a large dataset of arbitral awards will be challenging.

Second, assuming that arbitral awards are available, arbitral decision-making is not sufficiently high-volume to make it an ideal candidate for automation with AI. The annual number of international arbitration cases filed before the leading arbitral bodies is in the four-digits.[96] Thus, it is likely that the number of arbitral

---

are continuously working on designing one-shot learning methods where AIs can master a topic with fewer examples; Yaqing Wang et al., *Generalizing from a Few Examples: A Survey on Few-Shot Learning*, 53 ACM COMPUTING SURVEYS 3 (Mar. 29, 2020), https://doi.org/10.1145/3386252.

90. Katz et al., *supra* note 29; *See e.g.,* Ruger, et al., *supra* note 36, at 1150; N. Sivaranjani et al., *A Broad View of Automation in Legal Prediction Technology*, PROCEEDINGS OF THE THIRD INTERNATIONAL CONFERENCE ON ELECTRONICS COMMUNICATION AND AEROSPACE TECHNOLOGY [ICECA 2019], 2019, at 180, 182.

91. *See* Rhim & Park, *supra* note 80, at 30 (noting that obtaining the related materials in international arbitration is difficult, and thus, creating an AI that can predict arbitration result will be a struggle).

92. Albert Jan van den Berg, *Dissenting Opinions by Party-Appointed Arbitrators in Investment Arbitration*, *in* LOOKING TO THE FUTURE: ESSAYS ON INTERNATIONAL LAW IN HONOR OF W. MICHAEL REISMAN 821 n.4 (Mahnoush Arsanjani et al. eds., 2010) (noting that it is unusual to publish international commercial arbitration awards); Scherer, *supra* note 6, at 555; *See also* Marrow et al., *supra* note 6, at 68 (discussing the limited availability of arbitral awards in the context of U.S. domestic arbitration).

93. Del. Coal. for Open Gov't, Inc. v. Strine, 733 F.3d 510, 525 (3d Cir. 2013) ("Confidentiality is one of the primary reasons why litigants choose arbitration to resolve dispute."); Bette J. Roth et al., *1 Alternative Dispute Resolution Practice Guide* § 3:20 (2019) ("Business and professional persons consider the privacy of arbitration to be an advantage over litigation.").

94. *E.g.*, Klaus Peter Berger, *Cost Sanctions for Delaying Tactics in International Arbitration,* FINANCES IN INTERNATIONAL ARBITRATION LIBER AMICORUM PATRICIA SHAUGHNESSY 13, 13 (Sherlin Tung eds., 2020); Elina Zlatanska, *To Publish, or Not to Publish Arbitral Awards: That is the Question…*, 81 INT'L J. ARBITRATION, MEDIATION & DISP. MGMT 25, 25 (2015); Fulvio Fracassi, *Confidentiality and NAFTA Chapter 11 Arbitrations*, 2 CHI. J. INT'L L. 213, 221 (2001).

95. *E.g.*, Alexis Mourre, *The Case for The Publication of Arbitration Awards*, *in* THE RISE OF TRANSPARENCY IN INTERNATIONAL ARBITRATION 53, 72 (Alberto Malatesta & Rinaldo Sali eds., 2013) (urging that transparency should override the principle of confidentiality as far as the publication of arbitral awards concern); *Note to Parties And Arbitral Tribunals On The Conduct Of The Arbitration Under The ICC Rules Of Arbitration*, INT'L CHAMBER OF COMMERCE, para. 40-46 (Jan 1, 2019), https://iccwbo.org/content/uploads/sites/3/2017/03/icc-note-to-parties-and-arbitral-tribunals-on-the-conduct-of-arbitration.pdf (providing an opportunity for the potential publication of all ICC arbitration awards made as of 1 January 2019).

96. Markus Altenkirch & Malika Boussihmad, *International Arbitration Statistics 2018 – Another busy year for Arbitral Institutions* (Jul. 2, 2019), GLOBAL ARB. NEWS, https://globalarbitrationnews.com/international-arbitration-statistics-2018-another-busy-year-for-arbitral-institutions/ (compiling the numbers of cases, which are all in the four-digits, that the leading arbitral institutions

awards are also in the four-digits. Although there is no hard rule about the minimum number of observations in a data set, it is often insufficient to have only a few hundred to a few thousand observations.[97] Considering the overall estimated number of arbitration cases filed annually, it is questionable as to whether the sample size is large enough to produce a reliable result.

Finally, assuming that in an ideal world arbitral awards are fully available, and a few thousand awards would be large enough to create a data set, another problem arises from the variety of disputes that fall under arbitration. International arbitration frequently involves cross-border transactions that often reflect complex contractual relationships in specialized subject matters.[98] On the one hand, this is positive in terms of creating an AI arbitrator because variety allows AI to make a broader range of predictions, thus leading to a better job of matching the real world.[99] On the other hand, the variety of disputes might shrink the size of the applicable sample, thus reducing the number of observations overall.[100]

For example, assuming that the 4,000 or so awards rendered by the leading arbitration institutions are fully available and will be used to develop an AI arbitrator, these awards will more likely involve disputes from different areas of law. There might be 800 awards on commercial contracts, 700 awards on construction, 450 awards on energy, 200 awards on financial services and banking, and so on.

Any increase in the variety of disputes will correspondingly decrease the number of available observations in the dataset for each type of dispute. The input features that will be identified as having predictive value (e.g., type of plaintiff, type of defendant, type of contract, and the like) will more likely relate to the characteristics of cases since the arbitral tribunal acts as a trial court and adjudicates the dispute on the merits.[101] Hence, the entire corpus of 4,000 awards may not be used

received between the years of 2012 and 2018). *See also* Scherer, *supra* note 6, at 555 (stating that the number of arbitral awards in investor-state disputes is only in the double-digits per year).

97. Scherer, *supra* note 6, at 555; Lehr & Ohm, *supra* note 21, at 678-79 (". . . tens of thousands could be sufficient for some applications, and other applications operate optimally on hundreds of thousands or even millions of observations").

98. Paul Klaas, *International Commercial Arbitration*, CCARBITRATORS, 1-2 (2017), https://www.ccarbitrators.org/wp-content/uploads/International-Commercial-Arbitration.pdf ("Industries that are particularly active in international commercial arbitration include aviation and aerospace; energy; pharmaceuticals; banking; financial services; insurance; consumer products; mining; oil and gas; agriculture; construction and engineering; culture, media, and sports; health care; shipping; telecommunications; commodities; and professional services.").

99. The more variety is, the better the accuracy of the AI models trained on this data. Medvedeva et al., *supra* note 35, at 15 (underlying the variety of cases in the dataset and noting that the result might show lower but more realistic performance).

100. *See* Alozn & Galadari, *supra* note 5 (". . . for any given prediction method, the precision of litigation outcome prediction is as good as the availability of information related to comparable disputes and the corresponding decisions.").

101. *See id* (opining analogously in the context of construction disputes). Many researchers have already started to develop models for predicting court decisions, but these models have usually been deployed to predict international or supreme court decisions. *E.g.*, Lage-Freitas et al., *supra* note 50, at 1 (predicting decisions of the Brazilian Supreme Court); Katz et al., *supra* note 29 (predicting decisions of the U.S. Supreme Court); Octavia-Maria Sulea et al., *Predicting the Law Area and Decisions of French Supreme Court Cases* 1 (Aug. 4, 2017), https://arxiv.org/pdf/1708.01681.pdf (predicting decisions of the French Supreme Court); Aletras et al., *supra* note 29, at 3 (predicting decisions of the ECtHR). However, arbitration, by its nature, is similar to trial court litigation, and thus, the applicability of these studies remains limited. Therefore, the studies analyzing the lower court decisions should come to the forefront. Researchers have been working on developing models for predicting the outcome of construction litigation using lower court decisions. *E.g.*, Thaveeporn Pulket & David Arditi, *Construction litigation prediction system using ant colony optimization,* 27 CONSTRUCTION MANAGEMENT AND

for a specific prediction model. Since the narrower the task, the smarter the AI, it is reasonable to conclude that the first applications of AI to the field will be in specialized aspects of arbitration, like commercial disputes, energy disputes, and others.[102]

B. Generalization and Overfitting

A machine's ability to generalize the output function is fundamental to the success of a model. More specifically, an AI arbitrator will be expected to generalize the decision-making pattern that it saw and learned in training and predict the outcome of a novel dispute based on that training.[103] However, this ideal scenario may not occur in practice because of certain limits inherent in machine learning, one of which is called "overfitting."

Overfitting occurs when the algorithm learns the idiosyncratic features of the data, so much so that the machine starts to create patterns and rules that fit the data exactly, thereby failing to generalize when applied to other data sets.[104] Consider the following example: in 2019, researchers trained an AI to identify a variety of images, including a tench, a type of large fish.[105] Instead of learning how to identify the fish, the machine identified "fingers in front of a greenish background" because "tench images typically feature the fish [being] held up like a trophy, thus making the hand and fingers holding it a very predictive image feature."[106] Now consider this situation: what if an AI arbitrator learns to read document metadata, like the file size. It may then predict outcomes based on file size rather than the characteristics of the dispute.

Overfitting is a highly likely problem in the development of a machine arbitrator because AI is prone to overfitting if the number of input features is high and a small dataset is used for training.[107] As for the former, a machine arbitrator will likely have a high number of input features because the decisions will be more fact-driven and less focused on questions of law, as arbitrators act like a trial court rather than an appellate court.[108] In discussing AI arbitrators, Scherer reminded us of the

---

ECONOMICS 241, 241 (2009) (using the Illinois circuit court cases filed in the period 1987–2005); David Arditi et al., *Predicting the Outcome of Construction Litigation Using Neural Networks*, 13 COMPUTER-AIDED CIVIL AND INFRASTRUCTURE ENGINEERING 75, 81 (1998) (using circuit court cases to the extent retrieved from the appellate court records). These studies offer an analogous study method and imply that a model to predict the outcome decisions will take the characteristics of cases as input features. *See also* Chaphalkar et al., *supra* note 6, at 1827 (predicting the outcomes of construction arbitration).

102. *See supra* note 101 and accompanying text. This argument also in line with the AI that exists in our world today, also known as narrow AI. Harry Surden, *Artificial Intelligence and Law: An Overview*, 35 GA. ST. U. L. REV. 1305, 1337 (2019) (noting that today's AI systems excel in limited settings). *See* also *infra* Section 0.

103. ALPAYDIN, *supra* note 13, at 176.

104. STUART J. RUSSELL & PETER NORVIG, ARTIFICIAL INTELLIGENCE A MODERN APPROACH 705 (2016).

105. Wieland Brendel & Matthias Bethge, *Approximating CNNs With Bag-Of-Local-Features Models Works Surprisingly Well On ImageNet*, 4-5 (Mar. 20, 2019), https://arxiv.org/pdf/1904.00760.pdf.

106. *Id*. at 5.

107. RUSSELL & NORVIG, *supra* note 104, at 706; Michael A. Babyak, *What You See May Not Be What You Get: A Brief, Nontechnical Introduction to Overfitting in Regression-Type Models*, 66 PSYCHOSOMATIC MED. 411, 411 (2004).

108. Scherer, *supra* note 6, at 555. *See also* Arditi et al., *supra* note 101, at 81 ("In order to totally eliminate the possibility of overfitting it is recommended that future work experiment with reducing the large number of input . . . "). *See also* Pulket & Arditi, *supra* note 101, at 245 (identifying 39 different

complexity of topics in arbitration.[109] This lends support to the argument that AI arbitrators will require more input elements in dealing with multi-layered issues. As for the latter, it has been already noted that a lack of data is a problem arising from the division of the whole number of awards available due to the very nature of international arbitration—its confidentiality, there are a low number of arbitrations per year, and there are a wide variety of disputes that arise in arbitration.

Researchers often note that there is a trade-off between a complex model that is prone to overfitting and a simpler model that has low predictive accuracy.[110] Data scientists, therefore, need to take active measures to produce an "optimal balance between fit and complexity" to create an AI arbitrator that can predict an outcome with high accuracy without succumbing to overfitting.[111] This brings to the fore the importance of the human factor in developing and maintaining an AI arbitrator.

C. Bias in Training Data

Bias in training can interfere with machine learning in two different ways. First, if the training data is biased, the algorithm will simply reflect the existing bias by encoding and reproducing it.[112] An oft-cited, real-life example involves Amazon, a U.S. tech giant. Amazon's recruiting algorithm showed bias against women when it taught itself that male candidates were preferable and penalized women applicants because most resumes submitted to the company came from men, reflecting the tech industry's current male dominance.[113] Data, therefore, can simply reflect current societal or historical imbalances stemming from race, gender, and ideology producing outcomes that do not reflect true merit.

AI arbitrators can be disposed to show bias, for example, if previous awards reflect a pattern that is biased against consumers and in favor of companies.[114] The

---

attributes); David Arditi & Thaveeporn Pulket, *Predicting the Outcome of Construction Litigation Using Boosted Decision Trees*, 19 J. COMPUTING IN CIV. ENG'G 387, 390 (2005) (identifying more than 40 different input elements).

109. Scherer, *supra* note 6, at 556.

110. *E.g.*, ALPAYDIN, *supra* note 13, at 55; PEI WANG, THE LOGIC OF INTELLIGENCE 218 (2006); Edward K. Cheng, *A Practical Solution To The Reference Class Problem*, 109 COLUM. L. REV. 2081, 2093 (2009).

111. *See also* Zane Muller, *Algorithmic Harms to Workers in the Platform Economy: The Case of Uber*, 53 COLUM. J.L. & SOC. PROBS. 167, 205 (2020) (". . . overfitting is endemic to machine-learning algorithms and requires active measures by developers to overcome."); Matthew Hindman, *Perspectives on Computational Social Science: Building Better Models: Prediction, Replication, and Machine Learning in the Social Sciences*, 659 ANNALS 48, 54 (2015) ("Overfitting is treated as something to be estimated and managed, not something that must be avoided at all costs.").

112. *E.g.*, Jason R. Bent, *Is Algorithmic Affirmative Action Legal?*, 108 GEO. L.J. 803, 812 (2020); Bob Lambrechts, *May It Please the Algorithm*, 89 J. KAN. B. ASS'N, Jan 2020, at 36, 41; Cofone, *supra* note 88, at 1404; Joshua A. Kroll et al., *Accountable Algorithms*, 165 U. Pa. L. Rev. 633, 680 (2017); Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CAL. L. REV. 671, 680–81 (2016).

113. Jeffrey Dastin, *Amazon Scraps Secret AI Recruiting Tool That Showed Bias Against Women,* REUTERS, https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G (Oct. 9, 2018).

114. *See* Jean R. Sternlight, *Creeping Mandatory Arbitration: Is It Just?*, 57 STAN. L. REV. 1631, 1650 (2005) (". . . providers and arbitrators vehemently deny the charge that they are biased. . . . Yet, critics maintain that, consciously or unconsciously, arbitrators may slant the result in companies' favor."); Carrie Menkel-Meadow, *Do the ""Haves'' Come Out Ahead in Alternative Judicial Systems?: Repeat Players in ADR*, 15 OHIO ST. J. ON DISP. RESOL. 19, 53 (1999) (". . . we do not actually know much about whether one-shot consumers do worse in merchant operated arbitration. . . .We assume they do fare

same holds true in the context of investment arbitration, where an AI arbitrator may find a biased pattern and reach a conclusion preferable to investors at the expense of host states.[115] In discussing arbitrator bias, William Park reported on a case in which the arbitrator responded, "Italians are all liars in these cases and will say anything to suit their book" after one party cited a case involving Italians.[116] To take this implication to an extreme, what if the arbitrator had not been removed and the award had been allowed to stand? An AI arbitrator might use this data and teach itself that all Italians are liars, thereby refusing any evidence from Italians in the future.

Training data can also lead to biased output if the data is retrieved from a biased sample. For example, where a certain group is overrepresented or underrepresented. As early as 1662, the English scholar Graunt, who studied causes of mortality among the residents of London, found that data could be biased by social factors, such as relatives' unwillingness to report deaths from syphilis.[117] The data tainted by bias would fail to represent different groups in the outcome in correct proportions and distort any conclusions drawn from the data to be used for prediction purposes. Worse still, these predictions can be routed back as inputs, creating a vicious cycle.[118] This feedback-loop problem means that the latent threat of any bias risks becoming systematic over time.[119] If Public Health England, for example, were to rely on the data cited in this example to determine the allocation of its resources, it would potentially risk underserving individuals with syphilis.

In the international arbitration context, if an AI arbitrator trained on international sale-of-goods cases is fed a higher number of awards for sellers than buyers, it would predict the outcome of the case for the seller. To prevent this, it is essential to warrant a training sample that has an *equal number of cases* that can represent each party and each outcome.[120] In the ECtHR study, for example, the researchers curated a training sample that had an equal number of violation/non-violation

---

worse because we assume that dispute resolution systems chosen and maintained by one of the disputants therefore must benefit that disputant.").

115. *See* Rhim & Park, *supra* note 80, at 19 (noting that an algorithm bias in the AI system itself may create "a risk of reaching a conclusion preferable to a conglomerate or specific country in the process of arbitration."); Waibel & Wu, *supra* note 54, at 39 (finding that investment arbitrators are influenced by their policy views).

116. William Park, *Arbitrator Bias* 7 (Boston University School of Law, Working Paper No. 15-39, 2015) https://scholarship.law.bu.edu/faculty_scholarship/15 (citing In re The Owners of the Steamship Catalina and The Owners of the Motor Vessel Norma [1938] 61 Lloyd's Rep. 360 (Eng.)).

117. PRINCIPLES OF ROBOTICS & ARTIFICIAL INTELLIGENCE 234 (Donald R. Franceschett ed., 2018).

118. Cofone, *supra* note 88, at 1403.

119. *Id.*

120. *E.g.*, Qi Dong et al., *Imbalanced deep learning by minority class incremental rectification*, 41 IEEE TRANSACTIONS ON PATTERN ANALYSIS & MACH. INTEL. 1367 (2019) ("Most existing learning algorithms produce inductive bias (learning bias) towards the frequent (majority) classes if training data are not balanced, resulting in poor minority class recognition performance."); Medvedeva et al., *supra* note 35, at 8-9. ("Generally, the more data is available for the training phase, the better the program will perform. . . . If we blindly provide it with all the cases, it might only learn the distribution of `violation'/`non-violation' cases rather than more specific characteristics. For example, we might want to train a programme that predicts whether there is a violation of article 13, and feed it all 170 `non-violation' cases together with all 1230 `non-violation' cases. With such an imbalance in the number of cases per type, it is likely that the programme will learn that most of the cases have a violation and then simply predict `violation' for every new case (the performance will be quite high: 88% correct). In order to avoid this problem, we instead create a balanced dataset by including the same number of `violation` cases as the number of non-violation cases.").

cases.[121] This approach is also adopted in other studies for predicting court decisions.[122]

Nonetheless, most of the relevant studies have targeted higher court decisions with the intent of predicting whether a ruling stands or is reversed.[123] As stated earlier, since the arbitration process is more like a trial court litigation, the complex and diverse nature of the disputes falling under international arbitration may create a hurdle in curating a balanced sample. While cases where a seller has failed to perform its obligation under the Convention on the International Sale of Goods ("CISG") may be readily identified, undertaking such a binary classification task at scale for a diversified number of disputes poses a challenge.[124] Hence, this discussion supports the conclusion drawn in the previous subsection that the first applications of AI to the field will be in specialized aspects of arbitration.[125]

D. Feature Selection

Feature selection is where the data analysts decide which variables the AI model should observe and identify as most important when analyzing correlations and patterns for the prediction task.[126] As a simple example, if the task is to decide whether it is safe to drive through an intersection, traffic lights are informative and have predictive value, whereas the color of the car ahead is not.[127] However, identifying important features in most cases is not as straightforward as in the traffic light example. Identifying features from a data set that are relevant to a given prediction task requires significant forethought and skill, thus making it a form of art.[128] Hence, an analyst's choices may have enormous implications for the quality of the end-product algorithm.[129]

---

121. Aletras et al., *supra* note 29, at 8.

122. *E.g.*, Lage-Freitas et al., *supra* note 50, at 3; Medvedeva et al., *supra* note 35, at 9.

123. *E.g.*, Lage-Freitas et al., *supra* note 50, at 2 (predicting fully favorable decisions/ partially favorable decisions/ appeal denial decisions of the Brazilian Supreme Court); Medvedeva et al., *supra* note 35, at 5 (predicting the ECtHR case outcome based on violation/non-violation); Aletras et al., *supra* note 29, at 1 (predicting the ECtHR case outcome based on violation/non-violation).

124. *See* Scherer, *supra* note 6, at 553 (noting that complex and nuanced matters of facts and law can be difficult to classify into a binary model).

125. *See supra* Section 0 and notes 101 - 102 and accompanying text.

126. Jiliang Tang et al., *Feature selection for classification: A review, in* DATA CLASSIFICATION: ALGORITHMS AND APPLICATIONS at 37, 41 (2014) ("Feature selection . . . usually leads to better learning performance (*e.g.*, higher learning accuracy for classification), lower computational cost, and better model interpretability."); Bent, *supra* note 112, at 813; Lambrechts, *supra* note 112, at 42; Shannon Brown, *Peeking Inside The Black Box: A Preliminary Survey Of Technology Assisted Review (TAR) And Predictive Coding Algorithms For eDiscovery*, 21 SUFFOLK J. TRIAL & APP. ADV. 221, 253 (2016); Mikella Hurley & Julius Adebayo, *Credit Scoring in the Era of Big Data*, 18 YALE J.L. & TECH. 148 (2016) (". . . feature selection refers to the task of choosing a subset of input attributes that are most relevant to the problem and which have the greatest predictive potential.").

127. *E.g.*, James Franklin, *Feature Selection Methods for Solving the Reference Class Problem: Comment on Edward K. Cheng, "A Practical Solution to the Reference Class Problem,"* 110 COLUM. L. REV. SIDEBAR 12, 15 (2010).

128. Brown, *supra* note 126, at 253; James Franklin, *The Objective Bayesian Conceptualisation of Proof and Reference Class Problems*, 33 SYDNEY L. REV. 545, 560 (2011); Franklin, *supra* note 127, at 15. *See also* Josua Krause et al., *INFUSE; Interactive Feature Selection for Predictive Modeling of High Dimensional Dat*a, 20 IEEE TRANSACTIONS ON VISUALIZATION & COMPUTER GRAPHICS 1614, 1615 (2014) (describing the significance of feature selection in predictive modeling).

129. Aziz Z. Huq, *A Right to a Human Decision,* 106 VA. L. REV. 611, 675 (2020) (stating that human intentions necessarily guide the process of feature selection); Talia B. Gillis & Jann L. Spiess, *Big Data*

Just as feature selection is the art of machine prediction, law is the art of detail. For example, legislatures and courts commonly grapple with the use and interpretation of the words "and" and "or," as one small word may change the entire meaning of the text.[130] Similarly, in machine learning, the specifics required for making reasonably precise and fair determinations may lie in a degree of granularity and scope not met by the chosen features.[131] Thus, if the features that explain variation within a certain group are not incorporated, the model may be unable to distinguish among the members of the group.[132] In effect, this will lead to broad generalizations that may disadvantage and systematically discriminate against individual members of the group.[133]

Barocas and Selbst offered an example of the use of academic credentials in faculty hiring. If an algorithm assigns a particular numerical weight to the academic reputation of the school that the candidate attended, feature selection may systematically discriminate against candidates from more deprived backgrounds who less frequently attend prestigious schools for economic reasons, despite having comparable academic capabilities.[134] To complicate things further, the data available to the algorithm will never represent the unquantifiable complexities of the real world in their entirety because the data sets only contain what can be quantified and used in mathematical calculations.[135] Thus, even if feature selection is undertaken in the best possible way provided by the state of the art, there will still be a lacuna in terms of the number and relevance of the features that might have a decisive effect on the outcome.[136]

Feature selection will likely create unfair results in the context of international arbitration because a model will not typically include all possible relevant factors. As an example, in *Dow Chemical France v. Isover Saint Gobain* the ICC tribunal created a new doctrine known as the group of companies doctrine, and extended an arbitration agreement to non-signatory companies within the same group in the interest of fairness.[137] The Court did so because the non-signatory companies were cherry-picking by acting as a signatory when engaging in the negotiation,

---

and Discrimination, 86 U. CHI. L. REV. 459, 474 n.50 (2019) (noting that feature selection involves human discretion); Stephanie K. Glaberson, *Coding Over the Cracks: Predictive Analytics and Child Protection*, 46 FORDHAM URB. L.J. 307, 331 (2019) (noting that feature selection is one way where developers engage in models); Emily Berman, *A Government of Laws and Not of Machines,* 98 B.U. L. REV.1277, 1305 (2018) (stating that the subjective nature of feature selection makes it one of the places where human impact is significant); Andrew D. Selbst, *Disparate Impact in Big Data Policing*, 52 GA. L. REV. 109, 137 (2017) (noting that feature selection is unavoidably subjective); Lehr & Ohm, *supra* note 21, at 677 n.92; Barocas & Selbst, *supra* note 112, at 688.

130. *E.g.*, Ira P. Robbins, *"And/Or" and the Proper Use of Legal Language*, 77 MD. L. REV. 311 (2018).

131. Barocas & Selbst, *supra* note 112, at 688.

132. Pauline T. Kim, *Data-Driven Discrimination at Work*, 58 WM. & MARY L. REV. 857, 877 (2017); Barocas & Selbst, *supra* note 112, at 688.

133. Kim, *supra* note 132, at 877; Barocas & Selbst, *supra* note 112, at 688.

134. Barocas & Selbst, *supra* note 112, at 689.

135. Bent, *supra* note 112, at 813; Berman, *supra* note 129, at 1305.

136. Bent, *supra* note 112, at 813; Berman, *supra* note 129, at 1305-06.

137. Dow Chemical France, The Dow Chemical Co. and others v. ISOVER Saint Gobain, ICC Case No. 4131, Interim Award of 23 Sep. 1982, *reported in* 9 Y.B. COM. ARB. 131, 136 (1984) [hereinafter Dow Chemical].

implementation, and termination of the contracts but refusing to be bound by an arbitration agreement because it was not advantageous for their position.[138]

The data analyst can select the most known and relevant features in the data set to ensure the arbitration agreement bounds non-signatories, whether assignment, succession, or agency doctrine is involved.[139] Nonetheless, a given algorithm will not be able to consider the case in light of novel circumstances and create a new doctrine based on fundamental considerations of equity.[140]

E. Opacity Problems: AI as a Black Box

The call for transparency is at the center of the discourse on AI's implications in the legal context. As AI's impact on society grows, there is a consensus that the opacity of AI algorithms may pose a threat in identifying bias or preventing other potential harms from the outset.[141]

The first prong of the problem in achieving transparency stems from the technical aspects of AI algorithms. AI algorithms' decision-making processes typically operate in a black box that makes the algorithm opaque even to its designers, much less legal professionals and laypersons.[142] AI researchers note that some algorithms, especially artificial neural networks that mimic human brain neurons, are too complex for humans to comprehend and disentangle.[143] Thus, it is challenging to establish retroactively a causal nexus between a specific input and a specific

---

138. Gizem Halis Kasap, *Etching the Borders of Arbitration Agreement: The Group of Companies Doctrine in International Commercial Arbitration under the U.S. and Turkish Law*, 2 U. BOLOGNA L. REV. 87, 92 (2017).

139. *See* Chaphalkar et al., *supra* note 6, at 1831 (selecting contract provisions as features); Pulket & Arditi, *supra* note 101, at 245 (selecting features such as whether estoppel's doctrine involved or whether the provision of contract involved in their model predicting the outcome of construction litigation).

140. *See* Stavros Brekoulakis, *Rethinking Consent in International Commercial Arbitration: A General Theory for Non-signatories*, 8 J. INT'L DISP. SETT. 610, 616 (2017) ("Having a unique appreciation of the international settings of arbitration and a strong belief in the reformist power of international law, the *Dow Chemical* tribunal was able to conceptualize arbitration agreements in international business transactions as autonomous and independent of the main contract, and subject only to transnational legal rules.") *See also* Richard M. Re & Alicia Solow-Niederman, *Developing Artificially Intelligent Justice*, 22 STAN. TECH. L. REV. 242, 280 n.125 (2019) (explaining equity as preventing the rigid, rule-bound application of the law in unjust ways). *See also infra* Section 0.

141. *E.g.*, Berman, *supra* note 129, at 1321 (raising concerns arising from opacity in the context of the government's use of machine learning); Joshua P. Davis, *Law Without Mind: AI, Ethics, and Jurisprudence*, 55 CAL. W. L. REV. 165, 181–82 (2018) (examining the reasons for opacity and noting that even transparency may not enable humans to monitor AI); Tal Z. Zarsky, *Transparent Predictions*, 2013 U. ILL. L. REV. 1503, 1526-30 (2013). *See generally* FRANK PASQUALE, THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION (2016) (explaining concerns surrounding the use of machine-learning predictions).

142. *E.g.*, Jenna Burrell, *How the machine 'thinks': understanding opacity in machine learning algorithms*, BIG DATA & SOC., Jan.–June 2016, at 1, 10; Michael L. Rich, *Machine Learning, Automated Suspicion Algorithms, and the Fourth Amendment*, 164 U. PA. L. REV. 871, 886 (2016). (". . . machine learning tends to create models that are so complex that they become "black boxes," where even the original programmers of the algorithm have little idea exactly how or why the generated model creates accurate predictions.").

143. *E.g.*, Crootof, *supra* note 5, at 241; Scherer, *supra* note 6, at 562; Shlomit Yanisky-Ravid & Sean K. Hallisey, "*Equality and Privacy by Design": A New Model of Artificial Intelligence Data Transparency via Auditing, Certification, and Safe Harbor Regimes*, 46 FORDHAM URB. L.J. 428, 439 (2019); Rich*, supra* note 142, at 886; Barocas & Selbst, *supra* note 112, at 1907 (noting that even technical experts may not "weave a sensible story to account for the statistical relationships in the model"); Burrell, *supra* note 142, at 4-5.

output and vice versa.[144] This makes it hard to identify when an algorithm makes a mistake.[145]

The inability to disentangle the inner workings of an AI arbitrator's decision-making could undermine the legitimacy of AI arbitration and pose a risk that an award will be set aside or not enforced. The opacity of an AI arbitrator's decision-making may thwart the parties' ability to identify any grounds on which to mount a challenge. In the case of enforcement, inexplicability of an AI arbitrator's decision-making will lead courts to refuse confirmation of an AI arbitration decision. Lack of reasoning is already a basis for courts to refuse to enforce an arbitration award issued by humans, and that basis is not likely to change merely because AI arbitrators begin issuing awards.[146]

The next problem arises from the fact that AI algorithms are proper subjects of trade secret protection. Companies have legitimate interests in protecting their proprietary information and can refuse to disclose such information[147] due to concerns of privacy or the potential for gaming the system.[148] An AI arbitrator's algorithmic opacity, however, poses a prima facie challenge to the principles of procedural justice and the legitimacy of the process of international arbitration.

International arbitration has a pluralistic, diverse potential. The parties are free to interview and select arbitrators of their choosing, leading in theory to a diverse arbitration panel.[149] How would parties select their AI arbitrators? Should the parties have access to the selected features and their respective weights, or the right to audit the data fed into the algorithm before they choose?[150] Although these

144. Thomas Wischmeyer, *Artificial Intelligence and Transparency: Opening the Black Box,* REGULATING ARTIFICIAL INTELLIGENCE 76, 89 (Thomas Wischmeyer & Timo Rademacher eds., 2020); Crootof, *supra* note 5, at 241-42; Yanisky-Ravid & Hallisey, *supra* note 143, at 439; Burrell, *supra* note 142, at 1.

145. Achieving transparency and producing explanations for the outputs of black box models is an area of active research and led to the birth of a new wave called as explainable AI ("XAI"). It is, however, debatable whether XAI models will produce successful results both in terms of transparency and accurate predictability at the same time due to accuracy versus transparency trade-off. *E.g.*, Arun Rai, *Explainable AI: from black box to glass box*, 48 JOURNAL OF THE ACADEMY OF MARKETING SCIENCE 137 (2020); Cynthia Rudin, *Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead,* 1 NATURE MACHINE INTELLIGENCE 206 (2019); Berman, *supra* note 129, at 1316-17 (noting that "the more complex and powerful an algorithm, the more opaque it is likely to be" and to date, no XAI has been developed).

146. *E.g.*, Smart Systems Technologies Inc. v. Domotique Secant Inc., [2008] J.Q. No. 1782, 2008 QCCA 444. (Can. Que.) (finding that the arbitral tribunal's failure to provide reasons for its decision was a violation of public policy in Quebec and refusing to enforce the award).

147. Crootof, *supra* note 5, at 241-42; Burrell, *supra* note 142, at 3-4; Jessica M. Meyers, *Artificial Intelligence and Trade Secrets*, 11 LANDSLIDE (2019), https://www.americanbar.org/groups/intellectual_property_law/publications/landslide/2018-19/january-february/artificial-intelligence-trade-secrets-webinar/.

148. Wischmeyer, *supra* note 144, at 78; Kroll et al., *supra* note 112 at 685; Burrell, *supra* note 142, at 3-4.

149. Brekoulakis offers a diverse example of a panel, featuring an English barrister, who has graduated from Oxbridge, a Greek lawyer from a middle-class background and a South-American state officer. Brekoulakis, *supra* note 140, at 610.

150. *E.g.*, Wischmeyer, *supra* note 144, at 76 (reporting that the German Conference of Information Commissioners called for new laws to publish detailed information about how the algorithms works, such as disclosure of the classifiers and weights applied to the input data and the level of expertise of the system administrators and the like); Ric Simmons, *Quantifying Criminal Procedure: How to Unlock the Potential of Big Data in Our Criminal Justice System*, 2016 MICH. ST. L. REV. 947, 997-99 (2016) (arguing that to assess the algorithm's performance, all factors that the algorithm used and its historical accuracy must be transparent); Kate Crawford & Jason Schultz, *Big Data and Due Process: Toward a*

questions remain unanswered for now and require partnerships between legal scholars and other experts to be answered, they bring us to another significant issue—*technical illiteracy*.[151]

Technical illiteracy stems from the fact that the design of algorithms and the writing of code are specialized skills that are usually not shared by lawyers and judges. The syntax of programming languages differs from that of human languages.[152] Therefore, even if the propriety data is made available to the parties, their lawyers, or the courts, reverse engineering the algorithms[153] will still require legal persons to develop code and computational literacy, which may not be feasible in the short term.[154] Notwithstanding the problem of AI transparency, one can also question the transparency of the human decision-making system.

Since the beginning of the twentieth century, two jurisprudential schools of thought, "legal formalism" and "legal realism," have offered contrasting theories on judicial decision-making.[155] Today, many empirical studies have proved that judges engage in intuitive decision-making that reflects a variety of factors.[156] Thus, just as modern AI systems, the reasoning of human judges can be opaque, and act as a black box. Can we, then, conclude that we should adopt AI algorithms regardless of the absence of transparency since they operate much like human judicial decision-making as far as the black box issue is concerned?

Although we can accept that AI and human black boxes function alike, this is not sufficient to make them substitutes. Even if it is true that judges can first arrive

---

Framework to Redress Predictive Privacy Harms, 55 B.C. L. REV. 93, 117 (2014) (noting that "right to audit the data used to make the determination" is part of the due process).

151. Burrell, *supra* note 142, at 4.

152. *Id*.

153. ELDAD EILAM, REVERSING: SECRETS OF REVERSE ENGINEERING 3 (Robert Elliott et al., 2005) (defining reverse engineering as "the process of extracting the knowledge or design blueprints from anything man-made").

154. NICHOLAS DIAKOPOULOS, ALGORITHMIC ACCOUNTABILITY REPORTING: ON THE INVESTIGATION OF BLACK BOXES 26-30 (2013), https://academiccommons.columbia.edu/doi/10.7916/D8ZK5TW2

155. Legal formalists have claimed that law is determinate and that judges apply rules to the facts of a case rationally and deliberately, using the methods of deductive logic to produce a single correct result. *E.g.*, STEVEN J. BURTON, AN INTRODUCTION TO LAW AND LEGAL REASONING 3 (2d ed., 1995). For the pure formalist, legal decision-making is independent of any extraneous factors, and judges are highly skilled mechanics and work in a judicial system that is a "giant syllogism machine." *E.g.,* Burt Neuborne, *Of Sausage Factories and Syllogism Machines: Formalism, Realism, and Exclusionary Selection Techniques*, 67 N.Y.U. L. REV. 419, 421 (1992)("Pure formalists view the legal system as if it were a giant syllogism machine, with a determinate, externally-mandated legal rule supplying the major premise, and objectively 'true' pre-existing facts providing the minor premise. The judges' job is to act as a highly skilled mechanic."); Cesare Beccaria, *The interpretation of the laws*, *in* BECCARIA: 'ON CRIMES AND PUNISHMENTS' AND OTHER WRITINGS 14–16 (Richard Bellamy ed., Richard Davies tran., 1995). *See generally* Roscoe Pound, *Mechanical Jurisprudence*, 8 COLUM. L. REV. 605 (1908) (criticizing the notion of "mechanical jurisprudence"). Legal realists, on the other hand, have fiercely criticized this school and posited that law is indeterminate and filled with gaps and contradictions on many occasions. They argued that judges are human beings and thus have moods, political leanings, or emotions insomuch as that judges respond mainly to "the stimulus of the facts of the case" instead of legal rules and reasons. *E.g.*, Brian Leiter, *Rethinking Legal Realism: Toward a Naturalized Jurisprudence*, 76 TEXAS L. REV. 267 (1997); Jerome Frank, *What Courts Do in Fact*, 26 ILL. L. REV. 645, 656 (1932).

156. *E.g.,* Cass Sunstein et al, ARE JUDGES POLITICAL?: AN EMPIRICAL ANALYSIS OF THE FEDERAL JUDICIARY 20-23 (2006) (finding that the political affiliation of the appointing president, among others, considerably matters to judicial votes); Aletras et al., *supra* note 29, at 16 (arguing that their study results support "basic legal realist intuitions"). *See also* Chaphalkar et al., *supra* note 6, at 1828 (identifying "extrinsic factors" such as arbitrator's experience, technical expertise, cognitive skills, background characteristics, and human nature that can affect the arbitrator's decision-making process)

at a decision about a dispute through intuition, they still turn to statutes or caselaw to look for justifications for their decisions and apply objective judicial reasoning in their written judgments.[157] Leaving aside the debate as to whether judicial decisions are arrived at via objective reasoning, as the formalists suggest, or are merely an illustration of it, as the realists suggest, it is the reasoned decisions or awards themselves that enable parties, courts, and other stakeholders to initiate or conduct a judicial review of a decision or an award or to hold the decision-maker accountable.[158]

The legal framework of arbitration, just as in any legal institution, is designed with human decisionmakers in mind. Thus, for example, a court can refuse an unreasoned award and remand to the arbitrator for clarification, whereas we do not know whether and how such remand and clarification can be achieved with an AI arbitrator. We can, therefore, say that what we expect from AI in terms of transparency is the legitimacy, accountability, and fairness that human arbitrators achieve through the reasons provided in awards. Yet, it remains unanswered for now how AI arbitrators can produce a reasoned award or otherwise offset the need for one.

F. The Heart of the Matter: Emotional Intelligence and Arbitral Decision-Making

John McCarthy, the father of AI, once asked Joseph Weizenbaum, another AI pioneer, "What do judges know that we cannot tell a computer?"[159] McCarthy's answer was "nothing."[160] For him, AI can be built to make judicial decisions.[161] Conversely, Weizenbaum saw judicial decision-making as one of the tasks that AI should steer entirely clear of. Drawing from the example of an American judge who could not sit in a Japanese family court due to cultural differences between Japanese and American families, Weizenbaum argued that AI is alien to certain domains of thought and called into question the idea of computers replacing judges.[162] Weizenbaum's understanding of a judge was one who could show emotion and compassion, and offer counsel and empathy.[163] Do arbitrators also need to show emotion and compassion? More directly, what are the attributes of a human arbitrator that an AI arbitrator might not replace?

---

157. JEROME FRANK, LAW AND THE MODERN MIND 101 (1930) ("Judicial judgments, like other judgments, doubtless, in most cases are worked out backward from conclusion tentatively reached."); KARL LLEWELLYN, THE BRAMBLE BUSH: ON OUR LAW AND ITS STUDY 38 (1930) ("[W]ith a decision already made, the judge has sifted through these 'facts' again, and picked a few which he puts forward as essential - and whose legal bearing he then proceeds to expound".); Joseph C. Hutcheson Jr., *Judgment Intuitive The Function of the Hunch in Judicial Decision*, 14 CORNELL L. REV. 274, 287 (1929)

158. *See* Wischmeyer, *supra* note 144, at 78 (arguing that transparency in the context of AI should mean accountability and trust in the AI decision-making process)

159. JOSEPH WEIZENBAUM, COMPUTER POWER AND HUMAN REASON FROM JUDGMENT TO CALCULATION 207 (1976)

160. *Id.*

161. *Id*.

162. Weizenbaum argued that "computers can make judicial decisions," but "they ought not to be given such tasks" even if they may be able to make correct decisions. *Id*. at 227.

163. *See* MCCORDUCK, *supra* note 18, at 358. (reporting that, for Weizenbaum, "some sort of atrophy of the human spirit was taking place, an atrophy that trusted only 'science' to interpret reality")

The current state of AI technology is only capable of solving narrow problems.[164] Irrespective of how much it excels at solving these narrow problems, the problems are narrow. Thus, the technology is referred to as artificial narrow intelligence ("ANI").[165] There are, however, a high number of tasks that require artificial general intelligence ("AGI") technology, which does not yet exist.[166] Such technology, when developed, will be considered on par with human intelligence, because it is designed to perform any task that a human can.[167]

To decide whether arbitral decision-making requires ANI or AGI, one must first discern how human intelligence functions. Early psychologists treated intelligence as a single and complex entity.[168] Later observers have viewed human intelligence as involving multiple skills and abilities, and thus multiple intelligences that are independent of each other.[169] How human intelligence is formed remains a subject of ongoing debate between psychologists.[170] Nevertheless, Howard Gardner's long-celebrated theory of multiple intelligence provides a solid basis on which to discuss whether AI can or cannot reach the level of human intelligence.[171]

The theory of multiple intelligence proposes that human intelligence is comprised eight distinct and autonomous forms of intelligence that include interpersonal and intrapersonal forms.[172] Intrapersonal intelligence reflects an inward-looking approach, interreacting with the self, and the ability to understand one's own interests, motives, and, in general, one's own "feeling life."[173] In contrast, interpersonal intelligence reflects an outward approach, interacting with others, and the capacity to understand other people's desires, motives, and intentions.[174]

---

164. *E.g.*, Cassio Pennachin & Ben Goertzel, *Contemporary Approaches to Artificial General Intelligence*, ARTIFICIAL GENERAL INTELLIGENCE 1 (Ben Goertzel & Cassio Pennachin eds., 2007) (defining narrow AI as AI applications that demonstrate intelligence in one or another specialized area, such as chess-playing, medical diagnosis, or automobile driving); Ryan Calo, *Artificial Intelligence Policy: A Primer and Roadmap*, 51 U.C.D. L. REV. 399, 405 n.22 (2017) (stating that narrow AI is designed to solve a single problem). *See also* Jack M. Balkin, *The Path of Robotics Law*, 6 CAL. L. REV. CIRCUIT 45, 57 (2015) (calling AI as "special purpose human beings" that can substitute humans under certain circumstances and for specific purposes).

165. Pennachin & Goertzel, *supra* note 164, at 1.

166. *See e.g.*, Peter Voss, Essentials of General Intelligence: The Direct Path to Artificial General Intelligence, in ARTIFICIAL GENERAL INTELLIGENCE, *supra* note 164 at 131-33 (referring to human general learning ability when discussing AGI); Calo, *supra* note 164 at 405 n.22 (noting that AGI can accomplish more than one task without necessarily excelling in one).

167. *Id*.

168. *E.g.*, A. R. Jensen, *The psychometrics of intelligence*, *in* THE SCIENTIFIC STUDY OF HUMAN NATURE: TRIBUTE TO HANS J. EYSENCK AT EIGHTY 221, 223 (Helmuth Nyborg ed., 1997); H. J. Eysenck, *Introduction*, *in* A MODEL FOR INTELLIGENCE 1-10 (H. J. Eysenck ed., 1982).

169. *E.g.*, HOWARD GARDNER, FRAMES OF MIND: THE THEORY OF MULTIPLE INTELLIGENCES 1-13 (3rd ed., 2011); ROBERT STERNBERG, THE TRIARCHIC MIND: A NEW THEORY OF INTELLIGENCE 68 (1988)

170. It is, at least, certain today that AI has not reached the level of human intelligence regardless of it is one or many despite the debates on the psychology front. *See supra* note 166 and accompanying text.

171. GARDNER, *supra* note 169, at 294.

172. *Id.* at 77-151 (discussing linguistic, naturalist, logical-mathematical, musical, spatial, bodily-kinesthetic, interpersonal, and intrapersonal intelligences in detail). *See also* Katie Davis et al., *The Theory of Multiple Intelligences, in* THE CAMBRIDGE HANDBOOK OF INTELLIGENCE *supra* note 10, at 485 (discussing the theory of multiple intelligences that is first proposed by Howard Gardner).

173. GARDNER, *supra* note 169, at xxxvi (stating that intra-personal intelligence grows out of, and is organized around, the feeling life of the individual.).

174. *Id.* at 253; Davis et al., *supra* note 172, at 488.

Daniel Goleman employed the term "emotional intelligence" to encompass Gardner's interpersonal and intrapersonal concepts[175] and noted that emotional intelligence entails, among other things, empathy[176] and social skills,[177] buttressing his claim with a great deal of empirical evidence. Drawing on Gardner and Goleman's classification, an AI arbitrator cannot reach the level of human intelligence because of its lack of emotional intelligence.[178] As it stands currently, what AI exhibits is a high level of logical–mathematical intelligence.[179]

An AI arbitrator that can excel in logical–mathematical intelligence will process a numeric input and deploy a series of algorithms to generate an output. When doing so, however, the AI arbitrator will not be *aware* of the award that it will enter, nor will it *understand* why it has entered such an award. Since an AI arbitrator does not have a sense of self, it has no intrapersonal skills to gauge the success of its performance.[180] As Simmons succinctly puts it, the arbitrator may act as "an automated vacuum cleaner that sucks up the dust and the crickets without caring or even realizing the difference."[181] This may cause grave mistakes in arbitration, like the example of the AI algorithm that decided to sacrifice the life of the pilot in a flight simulation because it determined that crashing the plane was the optimal path to obtaining the highest landing score.[182]

---

175. DANIEL GOLEMAN, WORKING WITH EMOTIONAL INTELLIGENCE 24-25 (1998). Michael Beldoch coined the phrase emotional intelligence in a 1964 paper, but the term gained popularity in the mid-1990s with the Goleman's book. Micchael Beldoch, *Sensitivity to expression of emotional meaning in three modes of communication*, *in* THE COMMUNICATION OF EMOTIONAL MEANING 31, (J. R. Davitz et al., 1964). GOLEMAN, *supra* note 175. Major precursors of Goleman's work include John D. Mayer, who also studied emotional intelligence. *E.g.*, John D. Mayer et al., *Perceiving affective content in ambiguous visual stimuli: A component of emotional intelligence*, 54 J. PERSONALITY ASSESSMENT 772 (1990). *See also* John D. Mayer et al., *Emotional Intelligence, in* THE CAMBRIDGE HANDBOOK OF INTELLIGENCE, *supra* note 10, at 528-29 (surveying the literature on emotional intelligence).

176. GOLEMAN, *supra* note 175, at 318. (defining as empathy "[s]ensing what people are feeling, being able to take their perspective, and cultivating rapport and attunement with a broad diversity of people")

177. *Id.* (defining social skills as "[h]andling emotions in relationships well and accurately reading social situations and networks; interacting smoothly; using these skills to persuade and lead, negotiate and settle disputes, for cooperation and teamwork")

178. GARDNER, *supra* note 169, at 253. (It must be noted that machines can understand emotions to a certain extent, but they only read as data— basically a binary number. Even if we assume that they could learn how to recognize emotions to its full extent, they still would be conservatively tied to the reading instructions of its programmer. Hence, they will not, indeed, *feel* it.); *See* Marrow et al., *supra* note 6, at 38 ("Today humans determine the problems the computer is called upon to solve and humans define the instructions needed to solve those problems"); Dagmar Schuller & Björn W. Schuller, *The Age of Artificial Emotional Intelligence*, 51 COMPUTER 38, 41,45 (2018) (noting that to date, artificial emotional intelligence research has mainly focused on human emotion recognition and generation for conversational agents and robots, but "[a] claim for "real" emotion. . . needs a body and a physical connection to the real world.").

179. GARDNER, *supra* note 169, at 135 (Gardner defines this as an ability to develop equations and proofs, performing comparisons, and solve abstract problems, identify patterns.); *See also* Davis et al., *supra* note 172, at 488 (explaining Gardner's logical-mathematical intelligence).

180. Richard C. Waites & James E. Lawrence, *Psychological Dynamics in International Arbitration Advocacy*, THE ART OF ADVOCACY IN INTERNATIONAL ARBITRATION 69, 73 ("no arbitrators has ever made a decision without first taking into consideration their perceptions of the world around them and their previous life experience.").

181. Ric Simmons, *Big Data, Machine Judges, and the Legitimacy of the Criminal Justice System*, 52 U.C. DAVIS L. REV. 1067, 1095 (2018)

182. R. Feldt, *Generating Diverse Software Versions with Genetic Programming: An Experimental Study*, 145 IEE PROCEEDINGS - SOFTWARE ENGINEERING 228 (1998)

AI's lack of intrapersonal intelligence is also the reason that it can have only limited interpersonal intelligence.[183] While AI can interact with others on a certain level and answer questions based on the input provided by the user, like Siri or Alexa, it does not do so because it *understands* the question.[184] Instead, AI uses the data and instructions that it has learned during its training and applies it to a new input, which is merely logical–mathematical intelligence.

The legal scholarship on AI arbitrators mainly revolves around the capabilities of machine learning algorithms that can predict court decisions—less attention is given to the nature of the human arbitrator.[185] If we are to replace human arbitrators with non-human ones, taking the fullest account of human emotional intelligence merits consideration.

Human arbitrators do not apply the law purely as a matter of logic and legal syllogism or based on patterns and probabilities, which is how AI functions.[186] An arbitrator's personality, as well as their understanding of the world, is not only important for whatever decisions are rendered but also for the parties who appoint the arbitrators in the first place. Practitioners emphasize that arbitrators' ability to understand and interact with the parties can be reframed as "juridical open-mindedness," something that sharply delimits the success of the arbitration.[187] In line with this claim, a 2010 survey found that arbitration users attached importance to the soft skills[188] of arbitrators and consider these skills as having an impact on the efficiency and cost of a case.[189] It is thus not surprising that parties search for an arbitrator who has effective communicative skills and can show empathy, both of which require emotional intelligence. [190]

---

183. Adriana Braga & Robert K. Logan, *The Emperor of Strong AI Has No Clothes: Limits to Artificial Intelligence*, 8 INFORMATION 156, 13 (2017).

184. *See e.g.*, Simone Natale, *To believe in Siri: A critical analysis of AI voice assistants* 4 (Communicative Figurations, Working Paper No. 32, 2020) (arguing that "AI voice assistants activate an ambivalent relationship with users, giving them the illusions of control in their interactions with the assistant while at the same time withdrawing them from actual control over the computing systems that lie behind the interface") George Anders, *"Alexa, Understand Me"*, MIT TECH. REV. (Aug. 9, 2017) ("Humans can figure out that a friend who says "I haven't been to the gym in weeks" probably wants to talk about stress or self-esteem. For AI software, that's a hard leap. Sudden switches in topic—or oblique allusions—are tough, too.").

185. *See generally* Marrow et al., *supra* note 6; *see generally* Scherer, *supra* note 6.

186. *See* Tania Sourdin & Richard Cornes, *Do Judges Need to Be Human? The Implications of Technology for Responsive Judging*, *in* THE RESPONSIVE JUDGE INTERNATIONAL PERSPECTIVES at 87, 113 (2018) (opining that AI will not any time soon be able to replace the necessary and essential humanity of a human judge).

187. *See* Pierre Lalive, *On the Neutrality of the Arbitrator and of the Place of Arbitration*, RECUEIL DE TRAVAUX SUISSES SUR L'ARBITRAGE INTERNATIONAL 23, 27 (1984); *See* MARGARET L. MOSES, THE PRINCIPLES AND PRACTICE OF INTERNATIONAL COMMERCIAL ARBITRATION 120 (2008) ("A reputation for fairness, arbitrator integrity, and wisdom is a great asset to an arbitrator and also benefit the parties."); Edna Sussman, *Biases and Heuristics in Arbitrator Decision-Making: Reflections on How to Counteract or Play to Them, in* THE ROLES OF PSYCHOLOGY IN INTERNATIONAL ARBITRATION 45, 48 (Tony Cole ed., 2017) (discussing the non-legal factors that affect arbitrators' decision-making).

188. 2010 INTERNATIONAL ARBITRATION SURVEY: CHOICES IN INTERNATIONAL ARBITRATION 25, 28 (2010) http://www.arbitration.qmul.ac.uk/research/2010/. (The survey defines soft skills as "the ability to work well with the other members of the panel, the parties and their lawyers and generally adopt a helpful and friendly demeanor)[hereinafter 2010 INTERNATIONAL ARBITRATION SURVEY]; *See supra* note 176-177 (This definition is, in fact, analogous with Goleman's definition of social skills, which is one of the abilities forming emotional intelligence.)

189. 2010 INTERNATIONAL ARBITRATION SURVEY, *supra* note 188, at 25.

190. MOSES, *supra* note 187, at 136-37; Claudia T. Salomon, *Selecting An International Arbitrator: Five Factors To Consider*, 17 MEALEY'S INT'L ARB. REPORT 1, 2-3 (Oct. 2002); *See* Sundaresh

Moreover, international arbitration is not necessarily a fully adversarial process. Parties may agree to negotiation-based decision-making where the arbitrator uses his or her decision-making authority to facilitate resolution, which inevitably requires reading the feelings of the parties and handling those feelings adeptly.[191] This brings the importance of human emotional intelligence to the fore once more. In fact, even in an adversarial arbitration, arbitrators are often expected to handle proceedings in a manner bearing all the hallmarks of a proficient mediator.[192]

Overall, we do not know if "[b]y 2029, computers will have emotional intelligence and be [as] convincing as people" like the futurist Ray Kurzweil famously predicted.[193] Nonetheless, the question of whether AI arbitrators can substitute for humans requires us to consider human emotional intelligence thoroughly. This is because an AI arbitrator without emotional intelligence might not be as credible to the parties. Worse still, this may endanger the integrity of the arbitration process itself and the parties' faith in the overall process, or at the very least, lead to an uncertain adjudication between the parties. If the parties' resentment against arbitration becomes entrenched, they can become hesitant to relinquish their litigation rights, which subsequently harms the very characteristics of arbitration that make it so attractive in the first place.

V. ARBITRATION OF THE FUTURE BY AN AI ARBITRATOR

The effectiveness of any arbitration hinges on the courts' readiness to intervene and facilitate an ongoing or pending arbitration or, ultimately, enforce an award.[194] As such, arbitration agreements and awards, whether AI is involved or not, require the assistance of national courts and their respective laws. Given the ultimate relationship between the courts, national laws, and arbitration, this section explores

---

Menon, *Adjudicator, Advocate, or Something in Between? Coming to Terms with the Role of the Party-Appointed Arbitrator*, 34 J. INT'L. ARB. 348, 354 (2017) (stating that "used as a tool to overcome the distrust between disputants from diverse cultures" and"[w]ithout the comfort of being able to appoint an arbitrator of one's choosing, it was said that the parties could not be brought to the table."); Chiara Giorgetti, *The Arbitral Tribunal: Selection and Replacement of Arbitrators, in* LITIGATING INTERNATIONAL INVESTMENT DISPUTES: A PRACTITIONER'S GUIDE 143, 148 (Chiara Giorgetti eds.)(2014) ("Parties and their counsel spend substantial time and resources selecting the party-appointed arbitrator, which underline the importance of appointments."); Waibel & Wu, *supra* note 54 (Looking at the time spent on arbitration selection, Waibel and Wu also suggested that the personality and background of the arbitrator are decisive factors for arbitration outcomes.); *See* DANIEL SUSSKIND & RICHARD SUSSKIND, THE FUTURE OF THE PROFESSIONS: HOW TECHNOLOGY WILL TRANSFORM THE WORK OF HUMAN EXPERTS 251-54 (2015) (suggesting that the role and importance of empathy in the professions are often overstated).

191. Nathan Witkin, *Consensus Arbitration: A Negotiation-Based Decision-Making Process for Arbitrators*, *in* CONTEMPORARY ISSUES IN INTERNATIONAL ARBITRATION AND MEDIATION THE FORDHAM PAPERS 2011 434, 434-35 (Arthur W. Rovine ed., 2012) (discussing consensus arbitration in the context of international arbitration); GOLEMAN, *supra* note 175, at 24 (emphasizing the importance of empathy and social skills).

192. MATTI S. KURKELA ET AL., DUE PROCESS IN INTERNATIONAL COMMERCIAL ARBITRATION 149-50 (2nd ed., 2010).

193. SXSW News, *Interview by Douglass Cabellero with Ray Kurzweil, Director of Engineering, Google*, SXSW, at 5:30 PM (March 3, 2018) https://www.sxsw.com/news/2018/sxsw-live-streaming-schedule-2018/.

194. Dame Elizabeth Gloster, *Symbiosis or Sadomasochism? The relationship between the courts and arbitration*, 34 ARB. INT'L 321, 321-23 (2018); Stavros L. Brekoulakis, *International Arbitration Scholarship and The Concept of Arbitration Law*, 36 FORDHAM INT'L L.J. 745, 771–72 (2013).

some of the possible scenarios presented by an AI arbitrator as far as national arbitration laws and courts are concerned.

A. Can an Arbitrator Be an "It?" Appointing AI as an Arbitrator

A good starting point is to ask the most basic question, whether or not machines are eligible to sit as arbitrators under current arbitration legislation. In other words, are there express provisions concerning any requisite human qualities of arbitrators? The Convention on the Recognition and Enforcement of Arbitral Awards ("the New York Convention")[195] refers to arbitrators in two articles, Art.I (2) and Art.V (1)(b), but does not provide or imply that the arbitrators must be human beings.[196] The Convention simply refers to "the arbitrator" as an appointee that makes an award. That human qualifications are not specified is arguably not surprising considering that the Convention was ratified before the rise of AI, thus, the concept of an arbitrator was simply assumed to refer to a human being.[197]

Given that there is no explicit or implicit restriction against AI arbitrators in the wording of the New York Convention, AI arbitrators could technically issue an award that can be recognized or enforced under the Convention.[198] This is especially true if there is an explicit agreement between the parties to appoint an AI arbitrator.[199] Nonetheless, this will eventually depend on how courts come down on the issue and the extent to which they adopt a pro-arbitration approach.

Turning to the survey of national arbitration laws, it appears clear that AI's implications for arbitration—much less the question of whether a machine can be

---

195. 9 U.S.C.A. §§ 201–208 [1970] [hereinafter the New York Convention](The Convention on the Recognition and Enforcement of Arbitral Awards 21 U.S.T. 2617, 330 U.N.T.S. 3 9)(The Convention has been ratified by 161 countries to date); *Status: Convention on the Recognition and Enforcement of Foreign Arbitral Awards*, UNCITRAL, http://www.uncitral.org/uncitral/en/uncitral_texts/arbitration/NYConvention_status.html (last visited Aug. 1, 2020).

196. The New York Convention *supra* note 195 (Art. I(2) provides that "[t]he term 'arbitral awards' shall include not only awards made by arbitrators . . . but also those made by permanent arbitral bodies . . . ."Art.I(2) of the New York Convention *supra* note 195. The use of the term "arbitrator" only reappears at Article V(1)(b) in stipulating the recognition and enforcement of an award. But, the use in this clause does not have a bearing on the constitution of the term "arbitrator."); *See* The New York Convention *supra* note 195, at Art. V(1)(b) (providing that an arbitral award can be refused to be enforced if the "party against whom the award is invoked was not given proper notice of the appointment of the arbitrator"); GARY B. BORN, INTERNATIONAL COMMERCIAL ARBITRATION 293 (2nd ed., 2014)(Born argues that the New York Convention does not require arbitrators to be human beings. Thus, under Article II of the New York Convention, the Signatories requiring arbitrators to be a natural person might arguably violate their obligations to recognize arbitration agreements.)

197. Irene Ng (Huang Ying) & Valeria Benedetti del Rio, *When the Tribunal Is an Algorithm: Complexities of Enforcing Orders Determined by a Software under the New York Convention*, in 60 YEARS OF THE NEW YORK CONVENTION: KEY ISSUES AND FUTURE CHALLENGES at 121, 123-24 (Katia Fach Gomez & Ana M. Lopez-Rodriguez eds., 2019).

198. *Id.*

199. Gizem Halis Kasap, *AI in Arbitration: AI is Coming for Arbitrators, Too*, ELECTRONICALLY IN TOUCH (May 22, 2020, 12:00 AM), http://nysbar.com/blogs/EIT/2020/05/article-4-1.html; *See* Marrow et al., *supra* note 6, at 74 (emphasizing parties' freedom to modify arbitration rules through a mutual agreement); Rhim & Park, *supra* note 80, at 16 (underlining that parties' agreement is the key to arbitration procedure so that the parties can utilize AI as they see fit); David Allen Larson, *Arbitrator As Judge… And Judge of Jurisdiction Symposium: The End Of Arbitration As We Know It? Arbitration Under Attack*, 3 Y.B. ARB. & MEDIATION 93 (July 1, 2011) (arguing that arbitration is historically well suited to deploy technology when compared with courts).

an arbitrator—have not been subject to closer analysis by legislators. Nonetheless, such a survey yields an interesting result.

In the first group of national laws, the arbitrator is explicitly required to be a natural person with full capacity, as opposed to being a legal person.[200] The most notable standards are set by the French Code of Civil Procedure,[201] the Dutch Code of Civil Procedure,[202] and the Portuguese Voluntary Arbitration Law.[203] All of which stipulate that only "a natural person" or "individuals" can act as an arbitrator. In including outright restrictions, these laws take a strict position and make it clear that AI cannot act as an arbitrator.

The international arbitration laws in the second group do not explicitly state that arbitrators must be human beings but envisage them as such by requiring standards that only a human can meet or attribute certain characteristics that are exclusive to human beings. The arbitration laws in the People's Republic of China,[204] Indonesia,[205] North Korea,[206] and Vietnam,[207] require an arbitrator to have a certain number of years of experience as a judge or lawyer, or have specialized knowledge on the relevant topic, and so on. Similarly, the international arbitration laws of

200. *See* James Hope, *Can a Robot Be an Arbitrator?* STOCKHOLM ARBITRATION YEARBOOK 103, 111 (2019) (similarly observing that the discussion generally concerns the issue of whether legal entities can act as arbitrators).

201. Décret n°2011-48 du 13 janvier 2011–Art. 1450, Code de procédure civile, *translated in* The French Code Of Civil Procedure in English (2019) ("Only a *natural* person having full capacity to exercise his or her rights may act as an arbitrator."); Hope, *supra* note 200, at 108; Mohamed S. Abdel Wahab & Ethan Katsh, *Revolutionizing Technologies and the Use of Technology in International Arbitration: Innovation, Legitimacy, Prospects and Challenges*, *in* Arbitration in the Digital Age: The Brave New World of Arbitration at 27, 49 (Maud Piers & Christian Aschauer et al., 2018).

202. Artikel 1023 lid 4 RV. (Neth.), *translated in* DUTCH CIVIL LAW (DCL), http://www.dutchcivillaw.com/civilprocedureleg.htm (last visited Aug. 1, 2020) ("Any *natural* person with legal capacity may be appointed as arbitrator") (emphasis added).

203. The Portuguese Voluntary Arbitration Law, Law 63/2011 of 14 Dec. 2011 *published in* the Portuguese Republic's official journal, Diário da República, series 1—No 238, at 5276, available at Westlaw 9-518-9657, *translated in The new Law on Voluntary Arbitration - English Translation,* Portuguese Arbitration Association, https://arbitragem.pt/en/apa/projects-legislation (last visited Aug. 1, 2020) ("The arbitrators must be *individuals* and have full legal capacity."); Rui Manuel Moura Ramos, *The New Portuguese Arbitration Act (Law No. 63/2011 of 14 December on Voluntary Arbitration),* 16 Y.B. Priv. Int'l L. 25, 31 (2014-2015).

204. Zhonghua Renmin Gongheguo zhong cai fa (中华人民共和国仲裁法) [The arbitration law of the People's Republic of China] (promulgated by the Standing Comm. Nat'l People's Cong., Aug. 31, 1994, effective Sept. 1, 1995, rev. Sept. 1, 2017), art. 13, *translated in* http://www.cmac.org.cn/wp-content/uploads/2018/08/Arbitration-Law-of-the-Peoples-Republic-of-China-2017-Amendment.pdf (requiring experience as a lawyer, judge or academic and the like); Hope, *supra* note 200, at 109.

205. Arbitration and Alternative Dispute Resolution Act (Law No. 30/1999), art. 12 (Indon., *translated in* http://www.flevin.com/id/lgso/translations/Laws/Law%20No.%2030%20of%201999%20on%20Arbitration%20and%20Alternative%20Dispute%20Resolution%20(no%20elucidation).pdf (requiring arbitrators to be at least 35 years of age, have at least 15 years of experience in the field).

206. The Law on External Economic Arbitration (Decree No. 875/1999), art. 19 (N. Kor.), *translated in* https://www.international-arbitration-attorney.com/wp-content/uploads/2013/07/North-Korea-Arbitration-Law.pdf (requiring arbitrators to have experience as a lawyer or judge and the like).

207. The Law on Commercial Arbitration (Law No. 54-2010-QH12/2010), art 20 (Viet.), *translated in* https://eira.energycharter.org/component/attachments/attachments.html?id=5527&task=download (requiring arbitrators to possess a university degree and at least five years' of work experience in the discipline studied and the like); LE NET, ARBITRATION PROCEDURES AND PRACTICE IN VIETNAM: OVERVIEW, Practical Law 1-608-8005 (current through Apr. 1, 2015).

Egypt,[208] Finland,[209] Iceland,[210] Italy,[211] and Sweden[212] simply specify the capacity an arbitrator must exhibit. That is, has full capacity, or at least is not a minor, a bankrupt, or incapacitated. The necessary skills and abilities attributed to arbitrators under these laws suggest strongly that an arbitrator should be a human being. Nonetheless, given that second group laws do not explicitly require a natural person at the outset, it can be argued that they manifest legal lacunae, resulting in uncertainty and creating under-inclusive statutory contexts for AI arbitrators.[213]

The final group of laws offer the least-detailed stipulations as to the human requisites of arbitrators. Akin to the second group, these laws presume that the arbitrator is a natural person and thus refer to arbitrators using gender pronouns. What is different is that to act as an arbitrator, they do not bring qualifying standards that need to be met by a human. For example, the UNCITRAL Model Law, which many countries have adopted either verbatim or align with in spirit, refers to arbitrators using gender pronouns such as "him and "his."[214] Turning to the United States, Section 5 of the Federal Arbitration Act also address arbitrators using the pronouns "he" and "they."[215] The language in Section 26 of the English Arbitration Act of 1996 assumes that an arbitrator is mortal, as it stipulates that an arbitrator's authority ceases "on his death."[216] The use of such phrasing in the final group of laws suggests that an arbitrator is also assumed to be a natural person. When compared with the second group, however, these standards seem to be more lenient, as

208. Law No. 27 of 1994 (The Law Concerning Arbitration in Civil And Commercial Matters), *al-Jarīdah al-Ramīyah*, vol. 16 bis, 18 Apr. 1994, art. 16(1) (Egypt), *translated in* Egypt: Law No. 27 of 1994, 10 ARAB L. Q. 34, 39 (1995) (providing that arbitrators cannot be minor, bankrupt or subject to any incapacity or interdiction); Wahab & Katsh, *supra* note 201, at 49.

209. Arbitration Act (Act No. 967/1992), art. 8 (Fin.), *translated in* https://www.finlex.fi/en/laki/kaannokset/1992/en19920967_20150754.pdf (providing that arbitrators cannot be a bankrupt or incompetent).

210. Act on Contractual Arbitration (Act No. 53/1989), art. 6 (Ice.), *translated in* https://www.vi.is/files/act%20on%20contractual%20arbitration_849555187.pdf (requiring arbitrators to have full capacity).

211. Codice di procedura civile [C.p.c.] art. 812 (It.), *translated in* http://www.newyorkconvention.org/11165/web/files/document/1/6/16285.pdf. (providing that minors, incapacitated and mentally disabled persons, bankrupts, and persons barred from public office cannot be arbitrators); Hope, *supra* note 200, at 108.

212. LAG OM SKILJEFÖRFARAND [SWEDISH ARBITRATION ACT] (Svensk forfattningssamling [SFS] 1999:116) (Swed.), *translated in* https://sccinstitute.com/media/37089/the-swedish-arbitration-act.pdf.

213. *See* Lyria B. Moses, *Recurring Dilemmas: The Law's Race to Keep up with Technological Change*, 2007 U. ILL. J.L. TECH. & POL'Y 239, 250-54 (2007) (discussing the examples of legal uncertainty surrounding the introduction of new technology and the implications of such ambiguity).

214. U.N. COMM. ON INT'L TRADE LAW, UNCITRAL MODEL LAW ON INTERNATIONAL COMMERCIAL ARBITRATION, at 4, art 11(1), U.N. Doc. A/40/17, U.N. Sales No. E.08.V.4 (2006) ("No person shall be precluded by reason of *his* nationality from acting as an arbitrator" (emphasis added)) [hereinafter UNCITRAL Model Law]. The next article in UNCITRAL Model Law has a similar provision. *Id.* art 12(1) ("An arbitrator, from the time of *his* appointment…" (emphasis added)); *see also* Hope, *supra* note 200, at 104; Wahab & Katsh, *supra* note 201, at 49.

215. 9 U.S.C. § 5 (2021) (providing that "the court shall designate and appoint an arbitrator or arbitrators . . . as if *he* or *they* had been specifically named…" (emphasis added)). To demonstrate the reach of parties' freedom of contract, Baravati v. Josephthal, Lyon & Ross, Inc., 28 F.3d 704, 709 (7th Cir. 1994) (Posner, C.J.) ("Indeed, short of authorizing trial by battle or ordeal or, more doubtfully, by a panel of three monkeys, parties can stipulate to whatever procedures they want to govern the arbitration of their disputes; parties are as free to specify idiosyncratic terms of arbitration as they are to specify any other terms in their contract"); *see also* Wahab & Katsh, *supra* note 201, at 49.

216. Arbitration Act 1996 c. 23, § 23 (Eng.) http://www.legislation.gov.uk/ukpga/1996/23/data.pdf ("The authority of an arbitrator is personal and ceases on his death"); Hope, *supra* note 200, at 105; Wahab & Katsh, *supra* note 201, at 49.

they do not establish benchmarks. This lacuna opens the possibility for AI to act as an arbitrator in the future, given that there is no strict requirement that an arbitrator be a natural person.

Party autonomy is the linchpin of arbitration, and thus arbitration is a private process that the parties can structure as they see fit.[217] In the face of the law's struggle to keep up with technology, this principle is of crucial importance, at least when there are legal lacunae regarding who or what can act as an arbitrator.[218] Equally important is the approach of national courts to arbitration. Arbitration agreements and awards where AI is appointed as an arbitrator will likely be enforced in jurisdictions that go to great lengths to establish themselves as international arbitration hubs and whose courts generally defer to party autonomy.

B. Due Process Requirements

Regardless of what the applicable law is, there are certain minimum standards of procedural safeguards in arbitration—from a valid delegation of jurisdictional powers to arbitrators to enforcement of an award—that are essential.[219] These safeguards are said to constitute part of the "procedural magna carta of arbitration."[220] Traditionally, these safeguards have proven essential to maintain confidence in the integrity of arbitration. However, with the advent of AI arbitrators, one might expect that the parties, arbitral institutions, and courts will attach particular importance to due process safeguards until they build a certain level of trust.

1. Assessing the AI Arbitrator's Independence and Impartiality

It is generally accepted that an arbitrator must be both independent and impartial with respect to the parties and the dispute.[221] Though independence and impartiality go hand in hand, they have different, albeit complementary, meanings. The former concerns the arbitrator's potential financial, professional, or personal ties to a party or a dispute that might influence their decision, whereas the latter addresses any favoritism or prejudice shown by an arbitrator toward a party or the matter in dispute.[222] In this regard, the independence of an arbitrator can be ascertained more

217. *E.g.*, Volt Info. Scis., Inc. v. Bd. of Trs. of Leland Stanford Junior Univ., 489 U.S. 479 (1989); *see also* Thomas E. Carbonneau, *The Exercise of Contract Freedom in the Making of Arbitration Agreements*, 36 VAND. J. TRANSNAT'L L. 1189, 1190-91 (2003) (providing that "party agreement often provides the most significant rules for regulating arbitrations and conducting arbitral proceedings").

218. *See* Guillermo Argerich et al., *Could an Arbitral Award Rendered by AI Systems be Recognized or Enforced? Analysis from the Perspective of Public Policy*, KLUWER ARB. BLOG (Feb. 6, 2020), http://arbitrationblog.kluwerarbitration.com/2020/02/06/could-an-arbitral-award-rendered-by-ai-systems-be-recognized-or-enforced-analysis-from-the-perspective-of-public-policy/.

219. Hong-Lin Yu & Laurence Shore, *Independence, impartiality and immunity of arbitrators - US and English Perspectives*, 52 INT'L & COMP. L.Q. 935, 935-36 (2003).

220. JULIAN D. M. LEW ET AL., COMPARATIVE INTERNATIONAL COMMERCIAL ARBITRATION 95 (2003).

221. Park, *supra* note 116, at 6 (noting that "[t]he common assumption is that an arbitrator in international disputes must be both impartial and independent"); *see also* Bruno M. Bastida, *The Independence and Impartiality of Arbitrators in International Commercial Arbitration*, 6 REV. E-MERCATORIA 1, 2-3 (2007).

222. Yu & Shore, *supra* note 219, at 935-36; Doak Bishop & Lucy Reed, *Practical Guidelines for Interviewing, Selecting and Challenging Party-Appointed Arbitrators in International Commercial Arbitration*, 14 ARB. INT'L 395, 398-401 (1998).

or less objectively via external examination of their ties or relationships.[223] Impartiality, on the other hand, relates to the arbitrator's state of mind and is thus more subjective.[224] An AI arbitrator provides surprising implications in terms of independence and impartiality of an arbitrator.

a. The AI Arbitrator's Independence

Securing the independence of an AI arbitrator is less challenging compared to a human arbitrator. Thus, the problems arising from issues of arbitrator independence likely fall away once an AI arbitrator is deployed. First and foremost, no machine has sentimental relationships, relations of enmity, financial dealings, or links of group identification.[225] Unlike humans, AI arrives at a decision considering just the facts provided in the data—or, in other words, through logical–mathematical intelligence.[226] Given that AI lacks the emotional intelligence a human arbitrator presumably has, no AI arbitrator will ever be conflicted and will be free from external pressures when making decisions.[227]

Take the example of IBM's Ross, on the assumption that the platform has reached the point that it can now act as an arbitrator.[228] Arbitrator Ross would be independent even if one of the parties owns IBM stock or is employed by IBM because the algorithm follows the same rules for every decision it makes without regard to the party's affiliation, unless the algorithm is designed specifically to reward the IBM-affiliated party. This brings the importance of the design and programming of the algorithms to the fore once more.[229]

In traditional practice, an arbitrator has an ongoing duty to disclose their ties and any circumstances that may influence their judgment or create the appearance of partiality in the eyes of the parties.[230] With the use of an AI arbitrator, the focus will shift from anthropomorphic ties or relationships to how algorithms are programmed. Given that, a similar duty of disclosure can be identified concerning how an AI arbitrator is programmed to ensure that justice is done on an independent

---

223. *E.g.*, MOSES, *supra* note 187, at 140-41; Yu & Shore, *supra* note 219, at 936; Bastida, *supra* note 221, at 4.

224. *E.g.*, MOSES, *supra* note 187, at 140-41; Yu & Shore, *supra* note 219, at 936; Bastida, *supra* note 221, at 4.

225. *See* NICK BOSTROM, SUPERINTELLIGENCE: PATHS, DANGERS, STRATEGIES 29 (2014) ("There is no reason to expect a generic AI to be motivated by love or hate or pride or other such common human sentiments: these complex adaptations would require deliberate expensive effort to recreate in AIs. This is at once a big problem and a big opportunity"); Simmons, *supra* note 181, at 1081 ("Predictive algorithms hold great promise for increasing both actual and perceived neutrality of the decision-makers, because the algorithm follows the same rules in the same way for every decision it makes").

226. *See supra* Section 0.

227. *See* Crootof, *supra* note 5, at 236-37 (discussing the external factors in judicial decision-making); s*ee also* Thomas J. Buocz, *Artificial Intelligence in Court: Legitimacy Problems of AI Assistance in the Judiciary*, 2 RETSKRAFT - COPENHAGEN J. LEGAL STUD. 41, 44 (2018) (suggesting that the use of AI in the judiciary help to minimize the influence of extraneous factors).

228. *See* Karen Turner, *Meet 'Ross,' the newly hired legal robot*, WASH. POST (May 16, 2016, 6:00 AM), https://www.washingtonpost.com/news/innovations/wp/2016/05/16/meet-ross-the-newly-hired-legal-robot/ (noting that ROSS has been marketed as "the world's first artificially intelligent attorney").

229. *See generally supra* Sections 0-0.

230. KURKELA ET AL., *supra* note 192, at 120.

basis.[231] More specifically, arbitral institutions or parties should require that AI-arbitrator developers, in their disclosures, clearly describe how the AI arbitrator is programmed and whether certain features are scored in such a way that the independence of the machine arbitrator might be tainted.

Just as with human arbitrators, independence does not guarantee impartiality in AI arbitrators. For example, Northpointe corporation's Correctional Offender Management Profiling for Alternative Sanctions algorithm ("COMPAS") scores defendants to assist judges in deciding jail terms, sentencing, and probation.[232] COMPAS undertakes a risk assessment by using more than 100 factors, with the notable exclusion of race.[233] Among defendants who have identical risk scores, COMPAS reported the actual recidivism rate for black and white defendants as nearly identical.[234] Another way of saying this is that the algorithm is independent and neutral as to race.[235]

Nonetheless, a 2016 ProPublica study claimed racial bias in the COMPAS algorithm, finding that the COMPAS algorithm produces higher false-positive rates for black defendants than for white ones.[236] Though it is still debatable whether the COMPAS algorithm is racially biased, this example strongly indicates that a seemingly independent AI arbitrator could nonetheless be partial toward the parties or the dispute, and this brings us to the discussion of impartiality. [237]

b. The AI Arbitrator's Impartiality

An impartial arbitrator can be defined as one that is not biased in favor of, or prejudiced against, the parties or the case in dispute.[238] As stated earlier, independence and impartiality are interrelated concepts. In describing the distinction, Bishop and Reed stated that "[a]n arbitrator who is impartial but not wholly independent may be qualified, while an independent arbitrator who is not impartial must be

231. *See* James Ming Chen, *Models for Predicting Business Bankruptcies and Their Application to Banking and Financial Regulation*, 123 PENN ST. L. REV. 735, 749 ("Absent intentional discrimination in feature selection, however, it is hard to imagine how machine-based . . . decisions could support a disparate impact approach to liability."); *see generally also supra* Section 0.

232. *See generally COMPAS Risk & Need Assessment System Selected Questions Posed by Inquiring Agencies*, NORTHPOINTE, 2 (2012), http://www.northpointeinc.com/files/downloads/FAQ_Document.pdf.

233. Sam Corbett-Davies et al., *Algorithmic Decision Making and The Cost of Fairness*, 23 ACM SIGKDD INT'L CONF. ON KNOWLEDGE DISCOVERY & DATA MINING 797 (2017) (noting that such recidivism algorithms do not explicitly use race as an input) [hereinafter *The Cost of Fairness*]; Sam Corbett-Davies et al., *A Computer Program Used for Bail and Sentencing Decisions Was Labeled Biased Against Blacks. It's Actually Not That Clear*, WASH. POST (Oct.17, 2016), https://www.washingtonpost.com/news/monkey-cage/wp/2016/10/17/can-an-algorithm-be-racist-our-analysis-is-more-cautious-than-propublicas/ [hereinafter *Not That Clear*].

234. *The Cost of Fairness*, *supra* note 233, at 803; *Not That Clear*, *supra* note 233.

235. *See* Simmons, *supra* note 181, at 1082 (discussing COMPAS and noting that there is "strong evidence that the algorithm is 'neutral,' in that its results are identical across race.").

236. *See* Julia Angwin et al., *Machine Bias*, PROPUBLICA (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

237. *See* Simmons, *supra* note 181, at 1083 ("It is hard to say how . . . neutrality in results versus identical false positive rates translates into the procedural justice realm. In theory, neutrality in results should be the most important factor. . . . But the disparity in the false positive rate, combined with the fact that low income level or prior convictions (which are both correlated to race) increase a defendant's risk score, may understandably cause black and Latino defendants to perceive these algorithms as not "neutral" to their race.").

238. Bishop & Reed, *supra* note 222, at 398.

disqualified."[239] Such a way of thinking is also applicable to AI arbitrators. As in the COMPAS example, among many others,[240] an AI arbitrator that is programmed to be procedurally neutral might nonetheless yield biased results by reflecting any pre-existing bias in the training data, by using data that does not represent the real world accurately or because of feature selection or any other problem arising from its inputs and programming.[241] Thus, even if an AI arbitrator is substantively and procedurally neutral, its inputs and programming may taint the award with partiality. This could eventually lead to the setting aside of the award or denial of enforcement, given that a lack of impartiality is regarded as a sufficient basis under national arbitration laws and international treatises for doing so.[242]

The question then arises as to which standards the courts will apply in deciding whether an award made by an AI arbitrator is partial and whether such standards are suitable for determining an AI arbitrator's impartiality. Though there is no uniform standard, most jurisdictions adopt analogous tests to determine whether the arbitrator's impartiality is assured in actions to set aside or refuse the enforcement of an award. For example, under English law, a party needs to demonstrate that there was a "real danger of bias."[243] Under U.S. law, "evident partiality" is necessary to vacate an award, and Justice Black's opinion in *Commonwealth Coatings Corp. v. Cont'l Cas. Co.* specified that arbitral tribunals "must avoid even the appearance of bias."[244] Under the UNCITRAL Model Law and the laws based on it, the existence of "justifiable doubt" must be shown to set aside or refuse the enforcement of an award.[245]

Despite differences in these standards, finding a lack of impartiality on the part of the arbitrator in traditional practice depends significantly on the facts of each case and is difficult to prove, as courts require a high threshold.[246] Once AI arbitrators are in action, these standards may pose further challenges for the parties or

---

239. *Id*. at 400.

240. *E.g.*, James Vincent, *What a machine learning tool that turns Obama white can (and can't) tell us about AI bias*, THE VERGE (Jun 23, 2020), https://www.theverge.com/21298762/face-depixelizer-ai-machine-learning-tool-pulse-stylegan-obama-bias (finding that algorithm to scale up pixelated images often generate faces with Caucasian features, *e.g.*, turning a low-resolution picture of Barack Obama to a white man's picture); Henry Wong, *Siri, Alexa and unconscious bias: the case for designing fairer AI assistants*, DESIGN WEEK (Jan. 17, 2020), https://www.designweek.co.uk/issues/13-19-january-2020/unconscious-bias-ai-voice-assistants/ (discussing how voice assistants like Alexa or Siri struggle to recognize different accents); Nikhil Sonnad, *Google Translate's gender bias pairs "he" with "hardworking" and "she" with lazy, and other examples*, QUARTZ (Nov. 29, 2017), https://qz.com/1141122/google-translates-gender-bias-pairs-he-with-hardworking-and-she-with-lazy-and-other-examples/ (discussing the gender bias problem in Google Translate's algorithm); Gwen Sharp, *Nikon Camera Says Asians: People Are Always Blinking*, SOCIOLOGICAL IMAGES (May 29, 2009), https://thesocietypages.org/socimages/2009/05/29/nikon-camera-says-asians-are-always-blinking/ (Nikon cameras furnished with the blink detection identified many of the Asian users as blinking because the software that used Caucasians as its main data found that the Asian users' eyes were never open.).

241. *See e.g.*, CATHAY O'NEIL, WEAPONS OF MATH DESTRUCTION 21 (2016) ("Models are opinions embedded in mathematics."); Anupam Chander, *The Racist Algorithm?*, 115 MICH. L. REV. 1023, 1036 (2017) ("Even facially neutral algorithms will produce discriminatory results because they train and operate on the real world of pervasive discrimination."); *see generally supra* Section IV.

242. *E.g.*, Christian Borris & Rudolf Henneke, NEW YORK CONVENTION ON THE RECOGNITION AND ENFORCEMENT OF FOREIGN ARBITRAL AWARDS OF 10 JUNE 1958—COMMENTARY 329, 339 (Reinmar Wolff ed., 2012).

243. R v. Gough [1993] AC 646 (Eng.).

244. Commonwealth Coatings Corp. v. Cont'l Cas. Co., 393 U.S. 145, 150 (1968)

245. UNCITRAL Model Law *supra* note 214, Art.12 (1)

246. *E.g.,* KURKELA ET AL., *supra* note 192, at 204.

the courts to discover bias due to the problems arising from the transparency of AI. This is likely to increase the difficulty of proving bias on the part of the AI arbitrator. To begin with, the parties challenging the impartiality of an AI arbitrator might not be able to identify the problem because of the "black box" nature of AI algorithms mentioned above. Take the U.S. standard for example, while arbitral tribunals must avoid even the appearance of bias, the real concern in the case of AI arbitrators is, arguably, quite the opposite—no such "appearance" is possible in the first place.[247]

As stated earlier, it is not easy even for designers to detect when an AI arbitrator makes a mistake since the machine-learning-based programs define the rules on their own.[248] Lawyers, who often employ words to prove legal concepts, will likely struggle to understand algorithms in general, much less to identify whether a given algorithm has acted partially. It is therefore likely that the lawyers will need to work with experts to prove that the AI arbitrator has acted partially. However, the black box nature of AI is still a hurdle even for experts in detecting any biased rule or decision.

Another problem is the fact that commercial and government algorithms are themselves often proprietary. Returning to the COMPAS example, it is yet indiscernible whether COMPAS is biased because Northpointe has refused to disclose the details of its proprietary algorithm and how it works, "making it impossible to fully assess the extent to which it may be unfair."[249] In the case of human arbitrators, the duty of disclosure requires disclosure of any circumstances that might cast doubt on the arbitrator's impartiality before and throughout their appointment as arbitrator.[250] Moreover, if a human arbitrator acts in a biased way, he or she will be removed even if there is no evidence that a final decision would, in fact, be based on bias.[251] In contrast, identifying bias in AI algorithms is possible by examining the algorithm's results during or after the algorithm's use.[252]

As in traditional practice, arbitral institutions should be able to require the disclosure of facts or circumstances that may cast doubt on an AI arbitrator's

247. *See* Halis Kasap, *supra* note 199; *see also* Margot E. Kaminski, *Binary Governance: Lessons from the GDPR's Approach to Algorithmic Accountability*, 92 S. CAL. L. REV. 1529, 1539 (2019) ("Turning to algorithmic decision- making risks cloaking the very things we find problematic in human decision-making under a veneer of technical impartiality. And where human decision-making can often be contested, algorithmic decision-making . . . is often taken at face value and left unchallenged and unchallengeable.").

248. *See supra* Section 0.

249. Corbett-Davies et al., *Not That Clear*, *supra* note 233. In *State v. Loomis*, defendant Loomis argued that the proprietary nature of the algorithm prevented a challenge to its scientific validity and the data used and contended that COMPAS risk assessment at sentencing violated his due process rights. 881 N.W.2d 749, 753-54 (Wis. 2016). The Wisconsin Supreme Court ruled that such sentences cannot be challenged because the court used the algorithm only as a part of the decision-making process. *Id.* Although Loomis sought certiorari from the U.S. Supreme Court, the Court denied it in June 2017. Loomis v. Wisconsin, 137 S. Ct. 2290 (2017). New York City uses another predictive policing algorithm called Palantir. N.Y. Univ. v. N.Y.C. Police Dep't, 2017 N.Y. Misc. LEXIS 5138, at *2 (N.Y. Sup. Ct. Dec. 22, 2017). In a recent case concerning algorithm opacity, the petitioner invoked the public's significant interest in terms of the transparency of predictive policing and requested disclosure of the input data. *Id.* at *5. The Supreme Court of the State of New York rejected its request. *See supra* Section 0.

250. *E.g.,* KURKELA ET AL., *supra* note 192, at 120-22.

251. *See* Rhim & Park, *supra* note 80, at 7 (noting that such a biased perspective, nonetheless, calls arbitrator's independence into question in the eyes of the parties).

252. *See, e.g.,* SHANE, *supra* note 24, at 23; *see generally also* Ziad Obermeyer et al., *Dissecting racial bias in an algorithm used to manage the health of populations*, 366 SCIENCE 447 (2019).

independence and impartiality. This could be the AI arbitrator's algorithmic decision-making steps that may cause doubts as to its ability to serve impartially.[253] To prevent the disclosure of proprietary information, arbitration centers might outsource to third-party auditors to ensure all stakeholders that the system, as a whole, is not biased.

Nonetheless, an AI arbitrator acting as a black box may still produce biased results even if the companies are required to disclose certain details of the AI arbitrator's proprietary algorithm. Systematically auditing the AI arbitrator for bias can help identify some of these problems before an award is rendered. However, the ultimate goal should be to anticipate bias problems before they arise and design AI arbitrators to avoid them, given that AI has the significant potential to be impartial and offer greater objectivity in decision-making.[254] After all, in its current state, it is difficult to see why we would choose an inscrutable silicone-made AI arbitrator black box over a similarly inscrutable carbon-based, human one.[255]

2. Due Process Related to Facts

Like all forms of adjudication, arbitration consists of three elements: facts, laws, and application of the relevant rule of law to the facts of a case.[256] Remarkably, Kurkela and others referred to the application of laws to the facts as a form of intelligence, "referring to the intellectual exercise of reconstructing the past on the grounds of established facts and then applying the relevant laws to those facts."[257] What we aim to achieve with an AI arbitrator is a similar application, albeit artificially. There remains the question of facts and law bearing on whether we can replace a human arbitrator with an AI one.[258]

---

253. *See* NICHOLAS DIAKOPOLUOS, ALGORITHMIC ACCOUNTABILITY REPORTING: ON THE INVESTIGATION OF BLACK BOXES 28 (2013). ("This includes things like (1) the criteria used to prioritize, rank, emphasize, or editorialize things in the algorithm, including their definitions, operationalizations, and possibly even alternatives; (2) what data act as inputs to the algorithm— what it "pays attention" to, and what other parameters are used to initiate the algorithm; (3) the false positive and false negative rate of errors made in classification, including the rationale for how the balance point is set between those errors; (4) training data and its potential bias, including the evolution and dynamics of the algorithm as it learns from data; and (5) the definitions, operationalizations, or thresholds used by similarity or classification algorithms."); Wischmeyer, *supra* note 144, at 95 ("This might entail information on (1) the data basis of the system; (2) the models and the decision logic; (3) (data) quality standards implemented by the system operators; (4) the reference groups or profiles used by the system; (5) actual or potential inferences made by the system with regard to the individual concerned; etc.").

254. *See* Martin Petrin, *Corporate Management in the Age of AI*, 2019 COLUM. BUS. L. REV. 965, 1006 (2019) (opining that AI has the potential to designed to be completely unbiased that can lead to increased objectivity in decision-making); *see also* Anjanette H. Raymond & Scott J. Shackelford, *Technology, Ethics, And Access To Justice: Should An Algorithm Be Deciding Your Case?*, 35 MICH. J. INT'L L. 485, 522 (2014) ("In fact, a well-designed artificial intelligence algorithm could be bias free (at least to the extent that the programmers are also bias free), which is an advantage that cannot truly be guaranteed with human actors.").

255. *See* Volokh, *supra* note 5, at 1187 n. 154 (asking the same question in the context of judges).

256. KURKELA ET AL., *supra* note 192, at 141.

257. *Id*. at 142.

258. One might argue that an AI arbitrator will not necessarily follow the human arbitrator's way of resolving a dispute (*i.e.*, applying laws to the facts to decide the outcome) on the grounds that AI finds patterns hidden in its training data, even patterns its programmers did not expect, and thus, does not replicate the human way of thinking. Regardless, scientific studies tell us that facts and the law are still important because they have predictive value in reaching an accurate outcome. *E.g.*, Aletras et al., *supra* note 29 ("Our empirical analysis indicates that the formal facts of a case are the most important predictive

In an ideal world, an AI arbitrator would have access to clear facts and accurate statements about them. In reality, establishing the facts is not straightforward, as it seems in the abstract. The parties are expected to subjectively express their opinions on the nature of the facts that they deem relevant and support their legal position and their legal ramifications.[259] Even when the parties assess the facts in good faith as being without innocent misrepresentation or deliberate concealment, the arbitral tribunal is still faced with the daunting job of deciding which facts are indeed relevant and whether the evidence submitted is sufficient to prove the disputed fact.[260]

Equally difficult is deciding the question of law, which depends necessarily on the facts of the case. When looking at the question of law, any legal decision and rule could be described in a binary classification format, such as "whether (1) the tribunal has jurisdiction: yes/no; (2) the parties entered into a valid contract: yes/no; (3) one party breached the contract: yes/no."[261] However, similar challenges arise because the question of law cannot really be separated from the question of facts given that the law defines what the legally relevant facts are.[262] Overall, an AI arbitrator will need established facts and relevant law to decide the outcome.

Evaluating whether AI will be able to replace a human arbitrator invariably raises the question of how AI arbitrators will decide the facts. Thus far, the AI models used in predicting case outcomes have drawn conclusions from previously made judicial decisions and were thus based on the facts of the case articulated by the court.[263] Without suggesting any form of partiality or lack of neutrality on the part of the judges to justify the outcome, the facts stated in any judgment will reflect the selection of favorable facts or the disregard of unfavorable facts.[264] Therefore, it remains unanswered as to how an AI arbitrator will otherwise decide the facts.

If AI presents the parties' version of the facts, the AI arbitrator will need to decide which of the facts the parties are presenting are relevant, process the evidence, and decide whether the evidence submitted is sufficient to prove the facts. Ben-Ari and others observe that AI systems that are used as a tool in the different stages of legal proceedings constitute different pieces of the puzzle and can be put together to replace judges and rule over the world of law.[265] Drawing from this argument, for example, one might assume that a Siri-like AI system could question

---

factor"); Liu & Chen *supra* note 29 at 3 (indicating that the features extracted from the factual background of the case are one of the top three predictors concerning the predictive performance for all machine learning models).

259. KURKELA ET AL., *supra* note 192, at 143; Phillip Landolt, *Arbitrators' Initiatives to Obtain Factual and Legal Evidence*, 28 ARB. IN'L 173, 223 (2012).

260. KURKELA ET AL., *supra* note 192, at 145; Landolt, *supra* note 259, at 223.

261. Scherer, *supra* note 6, at 556.

262. KURKELA ET AL., *supra* note 192, at 141-42 ("For example, what is a reasonable time for a claim on faulty delivery? What is a faulty delivery? Facts and law work together.").

263. *E.g.*, Lage-Freitas et al., *supra* note 50, at 1 (predicting the decisions of the Brazilian Supreme Court based on the case descriptions, among others); Sulea et al., *supra* note 101, at 2 (predicting the decisions of the French Supreme Court based on the case descriptions, among others); Aletras et al., *supra* note 29 (predicting the decisions of the ECtHR based on the case circumstances, among others); Pulket & Arditi, *supra* note 101, at 241 (using the facts section of the Illinois circuit court cases filed in the period 1987–2005, among others).

264. *See supra* note 75-79.

265. Daniel Ben-Ari et al., *"Danger, Will Robinson"? Artificial Intelligence in the Practice of Law: An Analysis and Proof of Concept Experiment*, 23 RICH. J.L. & TECH. 3, 28 (2017). The authors state that these technologies should be applied in a holistic manner, but it remains unclear what constitutes a holistic manner. *Id.*

witnesses, and use other technology detecting emotion to evaluate the witness's response and detect whether they are lying,[266] thus eventually establishing—with the help of other AI technologies—the facts of the case. Then, these findings could be fed into an AI prediction model to predict the outcome of the case. In such a scenario, the case is decided in the final instance by an AI arbitrator consisting of multiple parts.

Human ability far exceeds machine learning, especially when it comes to commonsense knowledge and reasoning.[267] Humans can assemble disparate pieces of background knowledge and information through intuition, judgment, and imagination, which all play a crucial role in inference and problem-solving in general.[268]

It is a great irony that AI outperforms human computational ability in many aspects, and yet cannot process basic commonsense information that we expect other people, even children, to know and regard as self-evident.[269] For example, during a witness's direct testimony, the witness may instruct a Siri-like AI system to "call me an ambulance," to which the AI system might answer, "okay, from now on, I will call you 'an ambulance.'"[270] No human counterparts would likely give such an answer. This example alone shows that common sense is a critical component of interacting with humans and solving problems in a meaningful way.

In international arbitration, a tribunal may adopt either a passive or an active role in establishing the facts, but the tribunal usually has a lesser duty to establish the facts than the parties do.[271] Nonetheless, the arbitrators may need to intervene on their own account to fully understand the case, such as by asking questions of the parties, requesting documentary evidence be submitted, or calling witnesses.[272] As Kurkela and others aptly put it, "cases are won or lost on facts." In other words, establishing that the facts are correct is an essential element in securing legal protection for substantive rights.[273]

If the AI systems that parties use for arbitration are insufficiently capable of grasping the commonsense or scientific complexity of the world, they are unlikely to be able to decide whether there is a need to hear witnesses or experts or to take

266. *See* Jonas Gonzalez-Billandon et al., *Can a Robot Catch You Lying? A Machine Learning System to Detect Lies During Interactions*, 6 FRONTIERS IN ROBOTICS AND AI (July 31, 2019), https://www.frontiersin.org/articles/10.3389/frobt.2019.00064/full (discussing the availability of a lie detection algorithm based on machine learning).

267. Eyal Amir, *Reasoning and decision making*, *in* THE CAMBRIDGE HANDBOOK OF ARTIFICIAL INTELLIGENCE, *supra* note 12, at 191, 206 ("The terms "commonsense reasoning" and "commonsense knowledge" refer to a broad set of abilities that humans bring into their decision making and thinking.").

268. Danks, *supra* note 21, at 161.

269. Scott E. Fahlman, *Parallel Processing in Artificial Intelligence*, *in* PARALLEL COMPUTATION AND COMPUTERS FOR ARTIFICIAL INTELLIGENCE 3 (Janusz S. Kowalik ed., 1988).

270. *See* Will Knight, *A Tougher Turing Test Shows That Computers Still Have Virtually No Common Sense*, MIT TECH. REV., (July 14, 2016), https://www.technologyreview.com/s/601897/tougher-turing-test-exposes-chatbots-stupidity/ (discussing the lack of commonsense in AIs using the example of Siri's failure). *See also* Karen Hao, *AI still doesn't have the common sense to understand human language*, MIT TECH. REV., (Jan 31, 2020), https://www.technologyreview.com/2020/01/31/304844/ai-common-sense-reads-human-language-ai2/ (discussing the lack of commonsense knowledge in natural language models).

271. KURKELA ET AL., *supra* note 192, at 144-45; Landolt, *supra* note 259, at 223.

272. English Arbitration Act sec. 34; KURKELA ET AL., *supra* note 192, at 151; MOSES, *supra* note 187, at 176.

273. KURKELA ET AL., *supra* note 192, at 172.

any other appropriate action to establish the facts of the case.[274] This might ultimately violate due-process rights. Parties ought to be able to present their case and have the opportunity to be heard, otherwise the award runs the risk of being set aside or not enforced. After all, the legitimacy of arbitration depends on its fairness in the eyes of prospective parties and national courts.

3. The Duty to Give Reasons

An arbitral award is not a simple expression of an outcome. Instead, it is the place where and the instrument through which the tribunal justifies the outcome, mollifies the losing party by elaborating on why it lost, and reassures parties that they have been heard and recognized.[275] In fact, Landau was convincing in arguing that, in contrast to the actual decision contained within it, the nature and quality of the award itself dictates whether arbitration was a success or failure overall.[276] A reasoned arbitral award has the following four objectives.

First, through reasoning the tribunal lays out the basis for the award to one party rather than the other. Naturally, a naked decision stating that "X has won" without explaining the grounds on which the decision rests does not soothe the losing party.[277] Second, a reasoned award guards against any arbitrariness. Through reasoning, the adjudicative power exercised by the tribunal can be monitored by the parties, appellate arbitration tribunals, or the courts.[278] Third, and somewhat relatedly, either the court at the seat of arbitration or an enforcing court can review the award through its reasons and decide whether the award is subject to post-award scrutiny.[279] Lastly, awards have normative value, albeit less than that of court

---

274. *See* MOSES, *supra* note 187, at 137 (reporting that Pierre Mayer, who is a professor, arbitrator, and counsel in Paris, describes his perfect chair as one who has common sense, among others).

275. *E.g.*, Scherer, *supra* note 6, at 562; Hope, *supra* note 200, at 115. Historically, arbitration was used in resolving simple controversies between merchants where an immediate answer had been more important than the reasons itself. *E.g.*, Berger, *supra* note 94, at 16; Roger S. Haydock & Jennifer D. Henderson, *Arbitration and Judicial Civil Justice: An American Historical Review and a Proposal for a Private/Arbitral and Public/Judicial Partnership*, 2 PEPP. DISP. RESOL. L. J. ISS. 141, 145 (2002). In the modern international commercial arbitration practice, however, the requirement for a reasoned award is a norm that is prescribed in all institutional rules. MOSES, *supra* note 187, at 197. Today, not only the parties but also the courts attach more significance to the reasoned awards.

276. Toby Landau, *Reasons for Reasons: The Tribunal's Duty in Investor-State Arbitration, in* 50 YEARS OF THE NEW YORK CONVENTION: ICCA INTERNATIONAL ARBITRATION CONFERENCE 187 (Jan van den Berg ed., 2009).

277. Felix Dasser & Emmanuel O. Igbokwe, *Chapter III: The Award and the Courts, Efficient Drafting of the Arbitral Award: Traditional Ways Revisited – Lesson Learned from the Past?, in* AUSTRIAN YEARBOOK ON INTERNATIONAL ARBITRATION 2019, at 279, 284-85 (2019); Scherer, *supra* note 6, at 562; Strong, *supra* note 84, at 17; Margaret L. Moses, *Reasoned Decisions in Arbitrator Challenges*, 3 Y.B. INT'L ARB. 199, 199 (2013) ("Reasoned decisions . . . provide transparency and help parties understand how the process work"); Landau, *supra* note 276, at 187 ("It frequently requires the investment of very substantial time and funds by all parties. As the demands – and in particular the costs – of the process increase, so do the expectations of the parties as to the quality and detail of the award.").

278. *E.g.*, Strong, *supra* note 84, at 17, 19-20; Landau, *supra* note 276, at 189.

279. One may argue that most jurisdictions with the modern arbitration laws do not allow merit review. Thus, the need for a well-versed reasoned award does not equally apply. Nonetheless, international commercial arbitration case law indicates that the lack of (inadequacy of) reasons is ground for non-enforcement. *E.g.*, Soyak Int'l Constr. and Inv. Inc. v. Hochtief AG, Nytt Juridiskt Arkiv [NJA] [Supreme Court Reports] 2009_p. 12 T 4387-07 (Swed.) (finding that an award will be set aside if it lacks reasoning completely); Smart Sys. Tech. Inc. v. Domotique Secant Inc., 2008 CanLII 444 (Can.) (finding that the lack of reasons was contrary to public policy).

judgments, and, thus, third-parties or tribunals can benefit from the reasoning of the award, to either follow or depart from it when they face an analogous case in the future.[280]

An AI-arbitrator generated award raises three different concerns. First, current studies in predicting court decisions indicate that algorithms can predict the outcome with a high degree of accuracy but cannot provide reasons in the conventional sense.[281] While computer scientists, data analysts, and other knowledgeable specialists might understand why the algorithm decided that "X has won," legal actors and parties would face difficulties in apprehending the reasons underlying the outcome.[282] Besides, in some cases, AI models suffer from a lack of explicability and cannot provide reasons for a given prediction—the aforementioned black box problem.[283] As aptly remarked by Lehr and Ohm, it is epistemologically challenging to ask an algorithm to justify why each particular prediction has resulted.[284] In any case, it is unsatisfying in most legal contexts to explain a prediction generally by referencing the scientific steps taken by an algorithm to reach it.[285] In effect, courts may well set aside or refuse to enforce an award produced by an AI arbitrator on the grounds of insufficient or absent reasoning. Just as damaging, the legitimacy of the arbitral process may be diminished in the eyes of the parties.[286]

Second, even if AI algorithms can provide a reasoned award in the conventional sense,[287] reasons will lose their function as the mechanism through which

280. It should be noted that arbitration is customarily a private and confidential process. Arbitral awards, therefore, are usually not available to non-parties. Moreover, unlike common law, the doctrine of stare decisis does not apply in arbitration. For these reasons, the role arbitral awards as guidance for the future is strictly limited. Strong, *supra* note 84, at 15; Landau, *supra* note 276, at 192. Nonetheless, arbitral awards have still normative value as persuasive authority. Although this is also limited in commercial arbitration because the facts of a dispute are generally unique and unlikely to repeat themselves, the international commercial arbitration community is willing to follow the precedents to achieve predictability and consistency. Strong, *supra* note 84, at 15-6; Landau, *supra* note 276, at 192; *See also* Scherer, *supra* note 6, at 555-56 (noting the international commercial arbitration is usually dealing with distinct and varying issues).

281. *E.g.*, Lage-Freitas et al., *supra* note 50, at 2 (predicting whether the appeal is affirmed, partially affirmed, or denied); Katz et al., *supra* note 29 (predicting whether the judgment is affirmed or denied); Aletras et al., *supra* note 29, at 3 (predicting whether there is a human rights violation or not).

282. *E.g.,* Re & Solow-Niederman, *supra* note 140, at 275; Lehr & Ohm, *supra* note 21, at 708. *See also* Wischmeyer, *supra* note 144, at 87 (noting that algorithmic transparency requirements must not limit itself to providing information, but it also should enable those affected by the decision to react to a decision in a meaningful way such as by challenging it); Hannah Bloch-Wehba, *Access To Algorithms*, 88 FORDHAM L. REV. 1265, 1270 (2020) (opining that mere disclosure of internal workings of algorithmic decision-making tools is insufficient to vindicate accountability and transparency interests).

283. *E.g.,* Burrell*, supra* note 142, at 10; Rich*, supra* note 142, at 886. *See generally supra* Section 0.

284. Lehr & Ohm, *supra* note 21, at 708.

285. *Id*.

286. *See* Strong, *supra* note 84, at 17, 20 (opining that a well-reasoned award may lessen the risk of a judicial challenge by eliminating some grounds for non-enforcement).

287. One may query how AI technologies can achieve this task. For example, one can argue that text-generating AI could be used to produce reasoned awards or decisions. A neural network text generator trained on more than 82 million Amazon product reviews generate reviews that are, admittedly, very genuine. Alec Radford et al., *Learning to Generate Reviews and Discovering Sentiment* 6 (Apr. 5, 2017), https://arxiv.org/pdf/1704.01444.pdf. An AI correspondent prompted the model to write a positive product review specifying the product category, name, and the rating score, and the algorithms produced a comment that appears to be genuine. Kyle Wiggers, *OpenAI let us try its state-of-the-art NLP text generator*, VENTUREBEAT (Feb. 14, 2019 9:00 AM) https://venturebeat.com/2019/02/14/openai-let-us-generate-text-with-an-ai-model-that-achieves-state-of-the-art-performance-in-several-nlp-tasks/ ("We . . . supplied the Amazon model . . . for specifying the product category, product name, review score, and review summary — with this: *"A great phone."* With the product toggle set to "Cell Phones and

arbitrators bring novel doctrines into being. This is because AI cannot produce a new kind of outcome.[288] It can only mimic existing thought patterns or combine them to produce what appears to be a novel outcome.[289] This, however, will be only a mathematically different version of an already existing pattern.

In reality, arbitrators have far-reaching freedom to apply the law in a way that tailors the rule in light of the facts of each case and to interpret rules by adopting purposive analysis considering the international character of the dispute in order to avoid the shortcomings of applicable national law.[290] As Lord Mustill's put it, "[I]n making his award, the arbitrator . . . creates new rules, which he then applies retrospectively to the original bargain [and] in the absence of established norms, the arbitrator exercises a creative function, acting as a social engineer."[291]

Arbitrators are mighty and can exercise considerable discretion in determining and applying the law governing the dispute, but are expected to justify the use of their power through reasons.[292] This creative task can require developing a new pattern of thought that results in a different outcome than idiosyncratic national laws will produce.[293] An AI arbitrator, for example, will not be able to create a new doctrine as the human arbitrators did in *Dow Chemical* because it is incapable of stepping outside the box when exigencies and the demands of international commerce require this.[294]

Accessories" and the rating set to "5" (i.e., five stars out of five), it generated: "*The phone is great! I've had it since July of 2013. I will be purchasing a second one in the near future. Love the ability to text, play music and talk to others via the mic. Voice quality is great. The ear buds are a little uncomfortable after a few hours of wearing them. The only problem I've had are the battery life. I know, that's just a problem with technology, and it will go away after the phone is used for a while. Other than that, I love the phone. It has been the best phone for me. It's an excellent phone. The case is a nice touch, but the plastic part of the phone is the best part of the phone. It gets dirty easily. If you have a camera on your phone, you can use that phone for pictures. I've uploaded pictures of what I've shown my family what i've done since i got the phone.'*") AI technology is promising in fine-tuning such an algorithm and train with the previous reasoned awards to mimic reasoned awards.

288. *E.g.*, CHITRA LELE, ARTIFICIAL INTELLIGENCE MEETS AUGMENTED REALITY: REDEFINING REGULAR REALITY 13 (1st ed. 2019) (explaining that AIs cannot think outside the box because they cannot act differently from what they originally programmed to do); Scherer, *supra* note 6, at 557 (arguing that AI prediction models keep conservative approaches that are in line with previous cases); Stern, *supra* note 79, at 4 (noting that AI will accelerate reaching a specified goal but "will not be able to step outside of the system of primary rule generation to revise its overall goal"); Karen Maxwell, *Summoning the demon: robot arbitrators: arbitration and artificial intelligence,* PRACTICAL L. ARB. (Jan. 17, 2019), http://arbitrationblog.practicallaw.com/summoning-the-demon-robot-arbitrators-arbitration-and-artificial-intelligence/ ("Decision by robot must also be inherently conservative, with the associated risk of perpetuating trends and stifling development").

289. *E.g.*, LELE, *supra* note 288, at 13; Scherer, *supra* note 6, at 558; Stern, *supra* note 79, at 4; Maxwell, *supra* note 288.

290. Joanna Jemielniak. *Transnationalization of Domestic Law in International Commercial Arbitration Through Comparative Analysis: Challenges for Legal Profession*; 7 CONTEMP. ASIA ARB. J. 309, 320 (2014); Gabrielle Kaufmann-Kohler, *Arbitral Precedent: Dream, Necessity or Excuse?,* 23 ARB. INT'L 357, 364-65 (2007).

291. Michael Mustill, *The New Lew Mercatoria: The First Twenty-five Years*, *in* LIBER AMICORUM FOR THE RT. HON. LORD WILBERFORCE 149, 161 (Maarten Bos & Ian Brownlie eds., 1987).

292. Abul F.M. Manriruzzaman, *The Lex Mercatoria and International Contracts: A Challenge for International Commercial Arbitration?,* 14 AM. U. INT'L L. REV. 657, 693 (1999). *See also* JAN PAULSSON, THE IDEA OF ARBITRATION 16 (2013) (discussing arbitrators' great power and autonomy in deciding the disputes); Kaufmann-Kohler, *supra* note 290, at 364 (discussing arbitrators' broad discretion in determining and applying the law).

293. *See* Mustill, *supra* note 291, at 161 (noting the creative intelligence behind an arbitral award).

294. *See* Dow Chemical, *supra* note 137, at 136 ("The decisions of these tribunals [ICC arbitral tribunals] progressively create case law which should be taken into account, because it draws conclusions

Finally, and somewhat relatedly, an AI arbitrator's likely conservative approach will hamper the development of lex mercatoria.[295] Lex mercatoria is indeed a living and breathing subject that transcends the constraints of any given judicial tradition and keeps pace with the world of international commerce.[296] Given that an AI arbitrator is likely to adhere to pre-existing rules and past data in the algorithm, an award produced by an AI arbitrator may fall short in creating the openness and flexibility required to develop rules and standards of the lex mercatoria.

C. Public Policy

The public policy defense has often been invoked as if it were a blanket term in international arbitration, as the term is loose enough to allow for quite extensive coverage from due process violations to violations of substantive laws. Though it is not usually successful, it poses an insidious threat in international arbitration. Especially in countries that are less friendly to arbitration because their courts have been prone to interpret public policy more broadly.[297] In general, however, courts have tended to find public policy violations if the values underpinning state interests and sovereignty are trespassed, and especially where these values are so fundamental for the states concerned that any deviation becomes intolerable.[298] To that extent, an AI-rendered arbitral award could face certain questions, if not challenges, based on public policy violations.[299]

To begin with, national courts might find that any arbitration conducted by AI, as opposed to a human, would necessarily violate public policy. This is because judges have long praised themselves as flesh-and-blood persons, not robots that show no affection and merely recite laws.[300] Judges have long used the word "robot" as a rhetorical device to disparage *mechanical* legal reasoning and advance the kind of *human* legal reasoning purportedly ascribed to legal experts and not because the dispute before the court pertains to such technology.[301] Similarly, a Turkish court once asserted that judges should adjudicate the disputes before them "by applying a human touch" to the facts and the case in question, implying that justice should be dispensed only by human beings with the requisite concern, wisdom, and

---

from economic reality and conforms to the needs of international commerce, to which rules specific to international arbitration, themselves successively elaborated, should respond.").

295. Scherer, *supra* note 6, at 557.

296. *E.g.*, KURKELA ET AL., *supra* note 192, at 5; L. Yves Fortier, *New Trends in Governing Law: The New, New Lex Mercatoria, or, Back to the Future*, 16 ICSID REV. - FOREIGN INV. L. J., 10, 16-17 (2001); Mustill, *supra* note 291, at 151.

297. IBA SUBCOMMITTEE ON RECOGNITION AND ENFORCEMENT OF ARBITRAL AWARDS, REPORT ON THE PUBLIC POLICY EXCEPTION IN THE NEW YORK CONVENTION 5 (2015), https://www.ibanet.org/Document/Default.aspx?DocumentUid=C1AB4FF4-DA96-49D0-9AD0-AE20773AE07E [hereinafter REPORT ON THE PUBLIC POLICY EXCEPTION].

298. *See e.g.*, Ciments Français v. OAO Holding Company Siberian Cement, Highest Arbitrazh Court, Russian Federation, No. VAS-17458/11, 27 August 2012 (establishing that the award conflicted with a previous judgment of the courts that the forum violates public policy); Oberlandesgericht [OLG] [Higher Regional Court of Düsseldor] July 21, 2004, VI Sch (Kart) 1/02, 7 (noting that in order for the public policy exception to apply, violation of the fundamental principles of the legal, economic and social order of the state should be evident and sufficiently significant that the decision is unacceptable); REPORT ON THE PUBLIC POLICY EXCEPTION, *supra* note 297, 6-11.

299. Halis Kasap, *supra* note 199.

300. As Professor Ryan Calo indicated, judges have an increasingly outdated conceptual image of a robot. Ryan Calo, *Robots As Legal Metaphors*, 30 HARV. J.L. & TECH. 209, 218-19 (2016).

301. *Id.* at 210.

compassion.[302] Against this backdrop, courts might find that an AI arbitrator's lack of emotional intelligence violates public policy even if there is no outright provision as to whether arbitrators need to be human. This is particularly true if the court finds that adjudication by a human is one of the most fundamental values of the country concerned.[303]

Next, courts might find a violation of public policy if an AI arbitrator causes the arbitral procedure to suffer from serious irregularities. As stated earlier, the public policy exception operates as a blanket provision and, thus, a losing party tends to raise the defense of procedural unfairness both on due process and public policy grounds in order to strengthen their defense.[304] Therefore, the court may also set aside or refuse to enforce an award on the grounds of public policy if it finds an AI arbitrator lacks independence and impartiality, or an AI-made award fails to provide reasoning.

Finally, the sheer novelty of the concept of an AI arbitrator is itself a risk of a public policy violation due to fear of the unknown. Remarkably, the international arbitration community has shown a reluctance to adopt new technological innovations for fear that doing so may result in the setting aside or non-enforcement of an award.[305] This is not an entirely unfounded fear. For example, a national court found that an award violated public policy because the procedure the ICC Court had applied to review the award was unfamiliar to the national court and, thus, was unlikely to be well-understood generally.[306] Based on this assumed obtuseness, the court held that the procedure interfered with the arbitrator's independence, thereby violating public policy.[307] Therefore, in grappling with AI-rendered awards, courts may adopt a highly narrow and protective stance due to their unfamiliarity with the topic.

In conclusion, in considering a public policy violation, one must ask whether an AI-rendered award could run against the values entrenched in the concerned state's law and society. On the bright side, the application of the public policy defense is significantly narrower today than in the past and presents no significant

---

302. 1st Civil Chamber of the Turkish Court of Cassation, No: 1976/9370-13138, dated 31 Dec. 1976 (Kazancı İçtihat Bilgi Bankası) [Kazancı Case Law Database] (Turk.).

303. Ng & Benedetti del Rio, *supra* note 197, at 131; Argerich et al., *supra* note 218; *See also supra* Section 0.

304. *E.g.*, Albert Jan van den Berg, THE NEW YORK ARBITRATION CONVENTION OF 1958–TOWARDS A UNIFORM JUDICIAL INTERPRETATION 300 (1981); Inae Yang, *Procedural Public Policy Cases in International Commercial Arbitration*, 69 DISP. RESOLUTION J. 61, 65 (2014); Felix Dasser, *International Arbitration and Setting Aside Proceedings in Switzerland: A Statistical Analysis*, 25 ASA BUL. 444, 456 (2007).

305. *E.g.,* José María de la Jara et al., *Machine Arbitrator: Are We Ready?*, KLUWER ARB. BLOG (May. 4, 2017) http://arbitrationblog.kluwerarbitration.com/2017/05/04/machine-arbitrator-are-we-ready/?doing_wp_cron=1592168544.4940040111541748046875; Paul Cohen & Sophie Nappert, *The March of The Robots*, GLOBAL ARB. REVIEW (Feb. 15, 2017) https://globalarbitrationreview.com/article/1080951/the-march-of-the-robots; *see also* Berger, *supra* note 94, at 13. (noting that international commercial arbitration is currently under strict public scrutiny due to the legitimacy problems in investor-state arbitration, and AI poses a new challenge under these circumstances).

306. Mauro Rubino-Sammartoano, *The Keban Arbitration*, 45 J. CHARTERED INST. ARBITRATORS 211, 241 (1980).

307. Nuray Ekşi, *Yargıtay Kararları Işığında Icc Hakem Kararlarının Türkiye'de Tanınması ve Tenfizi [Recognition and Enforcement of ICC Arbitral Awards in Turkey in the light of the Court of Cassation Decisions]*, 67 ANKARA BAROSU DERGISI 54, 62 n.33 (2009) (Turk.) (It must be noted that the Keban Dam Arbitration became a rare instance in which the Turkish Court of Cassation overruled itself, though the Court's reasons to overrule itself is not clear); Rubino-Sammartoano, *supra* note 306, at 241.

obstacle in the way of enforcement of arbitral awards in general in pro-arbitration countries. Thus, the application of public policy grounds will depend on the reception of courts to technology.[308] Countries that are not able to keep pace with technological innovations or set aside or refuse to enforce an award on these grounds will likely see their positions deteriorate against the evolving nature of the international commercial arbitration market.

VI. CONCLUSION

Artificial intelligence has advanced to the point that machines can compare and contrast historical cases in order to predict the outcome of a dispute at hand, and AI is increasingly being deployed to do so. Against this backdrop, this article has asked explicitly whether AI will be able to replace human arbitrators, and what the legal implications of doing so will be. Of course, only the future can answer these questions definitively. However, this article offered a set of provisional conclusions, arguing that futurists are downplaying arbitration's complexity.

This article argues that the phenomenon of an AI arbitrator should not be downgraded to an AI application that is trained on historical cases to make accurate predictions ex-ante based on new dispute data. Even though the results of the most cited AI studies achieved over 70 percent accuracy in case prediction, these studies are heavily skewed toward appellate decisions. Therefore, future studies on arbitration cases that focus on cases where the court handed down an original decision on a dispute rather than acting as a higher court are necessary.

The limited data available regarding arbitral awards, AI's technical limitation, and AI's inability to embody emotions are all obstacles that may prevent the widespread use of AI arbitrators. Representing the real world's unquantifiable complexities in their entirety in the dataset continue to be a problem no matter how advanced technology becomes. Moreover, arbitration necessitates understanding parties' motives, struggles, and expectations from a given legal relationship, given that arbitration is a dynamic and multifaceted undertaking that typically demands an advanced level of emotional intelligence. Entering awards that most resemble those made by previous arbitrators in the past should not be the goal. The goal should be able to answer what the most just decision is, thereby preserving the role of equity in arbitration. Therefore, this article argues that it should always be essential for arbitration to have a human element attached to it.

The legal framework of arbitration is designed with human decisionmakers in mind. Deploying AI arbitrators without a feasible legal framework for the development, design, and application of artificial intelligence in arbitration could tarnish arbitration's reputation and eviscerate its real meaning, thus undermining its position and standing as a dispute resolution method. Hence, this article advocates for an in-depth analysis of the use of AI arbitrators and further research on AI's proper role in arbitration to prevent pre-mature deployment of AI arbitrators where they are no more than a technician of the law. This creates an opportunity to invite diverse stakeholders from technical and legal arenas to collaborate to build and deploy AI arbitrators in a way that ensures accountability and fairness and protects the

308. *See* Philippe Billiet & Filip Nordlund, *A new beginning – artificial intelligence and arbitration,* KOREAN ARB. REV., 26, 27-28 (2018) (arriving the similar conclusion for South Korean arbitration practice).

legitimacy of the arbitral process in the eyes of prospective parties and national courts.