DISPUTE RESOLUTION JOURNAL®

American Arbitration Association®
International Centre for Dispute Resolution®
AAA Mediation.org™

OCTOBER 2020 | VOLUME 74 | NUMBER 4

Artificial Intelligence and Arbitration: The Computer as an Arbitrator—Are We There Yet?

Paul Bennett Marrow* Mansi Karol† Steven Kuyan‡

> 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.
>
> *Baravati v. Josephthal, Lyon & Ross, Inc.*, 28 F.3d 704, 709 (7th Cir. 1994) (Posner, C.J.).

In this age where big data is commonplace and computers are becoming more powerful every day, Artificial Intelligence ("A.I.") has become a fact of life and is here to stay. Computer scientists posit that with enough data and properly designed algorithms, the "well-trained" computer should soon be able to produce an acceptable arbitration award. The necessary data set would hopefully include thousands of transcripts from actual arbitration proceedings; hundreds of thousands of actual awards; all known reported judicial opinions issued by courts throughout the United States embodying

*Paul Bennett Marrow is an attorney/arbitrator and a member of the American Arbitration Association's Commercial Panel. He teaches Domestic Arbitration at New York Law School. He can be reached at pbmarrow@optonline.net.

†Mansi Karol is the Director of ADR Services, Commercial Division at the American Arbitration Association in New York. She oversees administration of the large, commercial complex caseload, user outreach, and the panel of commercial neutrals in New York. She can be reached at KarolM@adr.org.

‡Steven Kuyan is the Director of Entrepreneurship at New York University's Tandon School of Engineering and Managing Director of the NYU Tandon Future Labs. He's an advocate for the responsible adoption of AI, founder of NYCai, adjunct faculty at NYU, as well as an investor and advisor to numerous startups. He can be reached at kuyan@nyu.edu.

the complete state of arbitration jurisprudence; all relevant statutes and rules used by judges, lawyers, arbitrators, and the administrators of the arbitration process; and all known journal and law review materials.[1] If these computer scientists are correct, and the authors of this paper believe they are, the results would unquestionably be a game-changer in resolving legal disputes in many parts of the third world as well as in many industrialized countries. Today, millions either have no access to an existing system of justice or have access only to find the system badly choked by bureaucratic inefficiencies, costs that are beyond their reach, and/or corruption. The disputants are left to either take the law into their own hands, often resulting in a violent confrontation, or abandon their claims altogether. Either way the result is a distaste for the system under which they live and a disregard for the rule of law.

Arbitration by computers without human intervention offers an exciting alternative. Computers have no schedules. They can handle numerous tasks simultaneously. Computers need only electricity and a connection to the internet in order to perform any assigned task. Once up and running, a computer-based system for arbitration would be relatively inexpensive and thus within the reach of most disputants, no matter their economic status. And most important, computers don't do personal favors, demand fidelity and/or take graft.

As will be discussed throughout this paper, A.I. provides the bridge connecting the disregarded and/or overlooked disputant to a meaningful disposition of a dispute. This will be welcomed news that in the end will serve to encourage a respect for arbitration as a meaningful way to resolve disputes based on the rule of law and available to anyone. While it is true that even with the advantages of A.I., the outcomes may never be absolutely perfect, still a determination by computer is far preferable to no determination at all.

Humans have routinely believed that they can always make improvements on the performance of computers, and this is reflected in the way

[1]No such data set currently exists and once development efforts get underway, due to the evolutionary pressures inherent to our systems of law, the results will continue to be a work in progress with no data set ever containing the full body of subject knowledge.

that computers have evolved over the last 100 years. Computer science is a never-ending process striving for perfection. Imperfections, once identified, are the subject of study and research and, more times than not, result in the discovery of a way to eliminate the imperfection. The demand for bigger and better never ceases. Breakthroughs such as the development of hardware capable of supporting "big data" and computers using quantum physics for the processing of complex algorithms and huge data sets confirm this. With time, the concept of arbitration by a computer will very likely become an acceptable norm, providing considerable advantages over courtroom-based dispute resolution.

Achieving this goal, however, will be challenging, and numerous questions will need to be addressed and answered. Here is a sampling:

1. Is there a risk that data can be biased, *i.e.* skewed by any number of conscious and/or unconscious factors; and if so, can bias be identified and eliminated?

2. Is the computer a "black box" operating in ways that are beyond our ability to understand, and if so, are there algorithms that can assist us in understanding how the computer is actually operating?

3. Assuming A.I. driven awards become a reality, will the award be persuasive, insightful, and timely, or mechanical, predictable, and rigidly correct?

4. Will state and federal laws that govern arbitration need to be amended to allow the use of A.I.?

5. Can a party consent to using a system that he/she doesn't fully understand?

6. Which disputes are best served by A.I. acting as an impartial neutral?

For the moment at least, the computer operates in a robotic manner, and this is likely to remain so for the foreseeable future.[2] Theorizing, creative

[2]Vincent C. Muller et al., *Future Progress in Artificial Intelligence: A Survey of Expert Opinion*, 376 Fundamental Issues of Artificial Intelligence 553-71 (2016).

thinking, and robust understanding, the things humans do best, remain beyond the reach of computer science. Today humans determine the problems the computer is called upon to solve and humans define the instructions needed to solve those problems. That is not to say that the computer is to be dismissed as a device good only for crunching numbers and/or performing other robotic-like tasks. Quite the contrary. The human brain isn't designed to absorb and process great quantities of materials used in advocacy. No single person could address, digest, process, and evaluate a data set composed of over one million emails, whereas a properly "trained" computer does this with ease. Human thought processes are slowed measurably in the face of large amounts of data. Efficiency and speed are the primary reasons humans invented the computer. A properly equipped and "well trained" computer can digest new data, look for patterns, and make predictions and recommendations. And it is not uncommon for a computer to spot an unknown trend or pattern.

Today's A.I. is used for tasks such as legal research, drafting of contracts, corporate records, preparation of research memos, drafting of pleadings, facilitating document discovery, and providing language translation and interpretation, to name only a few.[3] Computers can review existing documents, detect and report on deficiencies, and make recommendations on ways to improve what it has reviewed. Computers can review briefs before they are filed; update research; and eliminate grammar, spelling, and formatting errors. A computer can recommend variations of any argument and even propose new arguments that may not have been previously considered. Case management is incorporating A.I. for tasks such as scheduling of meetings, telephone answering services, docket control, and the creation and mailing of standard form letters. Support services for courts and arbitrators now incorporate A.I. A downside is that jobs associated with these tasks are being lost. The ranks of administrators, secretaries, law librarians,

[3]William S. Veatch, Artificial Intelligence and Legal Drafting, American Bar Ass'n Legal Analytics Committee Newsletter (Apr., 2019) https://www.americanbar.org/groups/business_law/publications/committee_newsletters/legal_analytics/2019/201904/ai-legal-drafting/.

and paralegals are thinning with each passing day as computers take over many duties humans formerly performed.

On the upside, there are examples of AI being used to handle routine, transactional matters. For example:

- Online Dispute Resolution ("ODR") is readily available at sites such as Modria. These services provide a structure for processing a dispute with access to a human mediator and tools for evaluating the merits of claims.

- The American Arbitration Association offers the Modria Resolution Center for certain kinds of disputes.[4]

- eBay has a dispute resolution center and offers advice on how to best use it.

These services and others like them stop short of offering a computer as an arbitrator.

Many states and local jurisdictions have begun to implement programs designed to allow citizens access to dispute resolution online:

- Michigan has a program allowing citizens to resolve certain civil disputes online.[5]

- In Ohio,[6] New York,[7] and Texas,[8] there are online programs allowing citizens to submit real estate taxation disputes and traffic ticket challenges, with humans making the final determinations.

---

[4]https://aaa-nynf.modria.com/.

[5]Michigan Courts, MI-Resolve, https://courts.michigan.gov/Administration/SCAO/OfficesPrograms/ODR/Documents/contact/index.html.

[6]Franklin County Municipal Court, Online Dispute Resolution, www.courtinnovations.com/ohfcmc.

[7]Modria Resolution Center, New York No-Fault Insurance, https://aaa-nynf.modria.com/.

[8]Texas Judicial Branch, eFile Texas Status (Aug. 29, 2017), http://www.txcourts.gov/media/1438816/efiletexas-status-jcit-20170829.pdf.

- Utah has a program for online resolution of small claim disputes.[9]

These efforts have begun to address the current shortfalls of the judicial system(s) in the United States and elsewhere,[10] suggesting that continued research will enhance the possibilities A.I. presents. For the moment, at least, perhaps the nomenclature "artificial intelligence" should be refashioned "augmented intelligence,"[11] *i.e.*, robotic intelligence to be used by humans to assist in reaching as accurate and unbiased a judgment as quickly as possible.

This article looks beyond "augmented intelligence" and asks the question: Can A.I. be trained to a level that allows A.I. to replace an arbitrator and make final, binding awards? From the perspective of arbitration as a process, the answer is, if not now, then soon. If it is assumed that within a given class of disputes, every conceivable outcome has already been determined and resolved, and all relevant data points were known, the task would be simple. Presented with a factual scenario and asked to compare it to identical cases labeled and described in a data set, the computer would easily make a prediction of the likely outcome. But the likelihood of the circumstances being identical for any two cases is low. In contested disputes, by definition the parties see the facts and the law from dramatically different vantage points, and resolution is often difficult for the human arbitrator. For the human, what is important are the facts, the evidence being produced to prove the facts and, to some extent, applicable law. The human arbitrator processes this information using deductive reasoning, logic, established rules and, when necessary, common sense. For the most part, human arbitrators are totally unaware of the thousands or even millions of similar cases and any statistically relevant patterns in those cases. That there is a data set containing similar cases with claimants winning 52% of the time

[9]Utah Online Dispute Resolution Steering Comm., Utah Online Dispute Resolution Pilot Project 3-4 (2017), https://ncsc.contentdm.oclc.org/digital/api/collection/adr/id/63/download.

[10]Vivi Tan, *Online Dispute Resolution for Small Civil Claims in Victoria: A New Paradigm in Civil Justice*, 24 Deakin L. Rev. 101 (2019).

[11]Jan W. Vasbinder et al., *Artificial or Augmented Intelligence? The Ethical and Societal Implications*, *in Grand Challenges for Science in the 21st Century*, 51-68 (Jan W Vasbinder et al. eds. 2018).

and respondents 48% of the time is of no interest and has no influence on the arbitrator's decision. On the other hand, computer science instructs computers to make determinations using only pattern recognition and statistical formulations, with no consideration given to deductive reasoning, logic, or common sense. It's not surprising that computers and humans may literally see and process identical facts differently. The task ahead is figuring out how to unify these perspectives, so it is faithful to our systems of law *and* allows for an analysis tailored to the individual case at hand.

Today, there is no shortage of unresolved disputes of all kinds: large and small, simple and complex. And there is nothing suggesting that in the foreseeable future the number will not grow, perhaps even at an exponential pace. Driven in part by judicial and legislative findings that virtually any dispute is arbitrable, arbitration has become a major go-to alternative to the courthouse. Unfortunately, arbitration, once championed because of cost savings and other efficiencies, has become just the opposite, placing arbitration beyond the reach of many who seek an efficient, cost effective, disciplined, and fair alternative to the courthouse. Decision-making machines can address these deficiencies, making arbitration appealing.[12]

As the reader will see shortly, not every case is likely to lend itself to arbitration using A.I. Disputants and their lawyers involved in what have been called "bet the company cases" are the least likely to want to entrust such issues to a computer. In all likelihood, the best matches will prove to be small claim disputes stalled by a clogged judicial system that is bureaucratic, overwhelmed, and/or encrusted with corruption, and cases involving similar facts. In addition, remote access to a decision-making machine using internet technologies should eliminate bureaucratic delays and increase the ability of many who lack the means, or the time needed to travel distances to a courthouse, to resolve their disputes. The inability to obtain an unbiased and timely determination of a dispute creates a real risk that the rule of law will break down with parties even taking the law into their own hands. Recent history has shown us that one characteristic

[12]Omar-Rabinovich-Einy & Ethan Katsh, *Access to Digital Justice: Fair and Efficient Processes for the Modern Age*, 18 Cardozo J. Conflict Resol. 637 (2017).

of a failed state is widespread frustration across a population because of a lack of access to a practical, timely, and fair process for resolving matters important to them. No matter the type of case, it is reasonable and in fact, necessary, to see if A.I. can address the challenges presented. Only time will tell.

Acceptance of an award fashioned by A.I. requires that parties trust what the computer appears to be doing, *i.e.*, processing information and reaching an appropriate determination. Realistically though, the public's current perception is that the computer is a "black box," whose operations are beyond comprehension, principally because the computer cannot account for how and why it reached a given outcome. This lack of transparency has resulted in demands for a so called "right to explanation." Addressing this demand has generated extensive research into the development of appropriate procedures, including algorithms capable of compelling accountability and transparency, and has triggered governmental responses and demands:

- In 2016 the European Union's Parliament adopted the General Data Protection Regulation ("GDPR")[13] that became effective in 2018.
- In late 2017, New York City, suspicious of algorithms used to determine the allocation of everything from food stamps to firehouses created a fact-finding task force to determine if the algorithms were performing in a fair and equitable manner.[14] The task force released its report in November, 2019.[15]

I. Arbitration as the ideal candidate for computer-driven dispute resolution.

The first step is finding out if A.I. and arbitration are indeed compatible. The academic literature to date has had little to say about this relationship.

[13]Regulation 2016/679, GDPR art. 12, 2016 O.J. (L 119) 39-40 (EU) (addressing "Transparent information, communications and modalities for the exercise of the rights of the data subject").

[14]Julia Poweles, *New York City's Bold, Flawed Attempt to Make Algorithms Accountable*, The New Yorker (Dec. 20, 2017), https://www.newyorker.com/tech/annals-of-technology/new-york-citys-bold-flawed-attempt-to-make-algorithms-accountable.

[15]New York City, Automated Decision Systems Task Force Report (Nov. 2019), https://www1.nyc.gov/assets/adstaskforce/downloads/pdf/ADS-Report-11192019.pdf.

The focus of research to date has been on the relationship between A.I. and courthouse litigation. Arbitration is quite different from courthouse litigation. Arbitration is not intended to be a *substitute* for litigation. Arbitration is a meaningful *alternative*. This article argues this difference is so consequential as to make arbitration a far better candidate for A.I. applications than litigation. Arbitrators and A.I. are subject to constraints on how each operates, constraints not imposed on a judicial system. For A.I., performance is limited by the quality and quantity of training data and the quality and robustness of the algorithms. For the human arbitrator, performance is limited by the terms of the arbitration clause and governing law.

First consider the limitations placed on the arbitrator. Not being judges, arbitrators are rarely, if ever, granted authority to consider factors external to the case at hand; factors such as political trends and changes in societal norms. Arbitrators cannot disregard, modify, or defang existing law and/or precedent. Arbitrators can do no more than consider a prescribed factual scenario and apply the law as required by the parties. The arbitrator's first obligation is to the parties and to the terms of the agreement to arbitrate. By contrast, judges have a first obligation to the law without concern for any agreement by parties limiting the ability to apply, interpret, and even nullify statutes, precedents, and rules and regulations. Judges are free to consider changing societal norms and conditions. Judges can resolve a case of first impression, *i.e.*, a case not known to have been evaluated by any judge, and proclaim rationales for overturning precedent or voiding statutes and administrative rules and regulations.[16]

Next consider limitations placed on A.I. Unless allowed by the designer, A.I. can't operate outside its instructions. A.I. must obey the mathemati-

[16] An arbitrator's merit-based decisions are for the most part beyond judicial review. This is true even if the arbitrator makes a good faith mistake. Judicial review is limited to the vacating of an award on limited grounds specified by applicable federal and state statutes. *Major League Baseball Players Ass'n v. Garvey,* 532 U.S. 504, 509 (2001) The rule is followed by most state courts. *See, e.g.*, *Heimlich v. Shivji*, 7 Cal. 5th 350, 367 ("A court's power to correct or vacate an erroneous arbitration award is closely circumscribed."); *In re Santer*, 23 N.Y. 3d 251, 263 (2014) Courts, on the other hand are subject to review based on errors.

cal and structural limitations humans impose. The human designer restricts A.I.'s understanding of our world to the training data provided. Unless instructed to do so, A.I. is unable to access data on its own initiative. A.I. doesn't even know there is a world beyond what humans define for it.[17] While arbitrators can think and reason, they are constrained by the legal limitations of the process. A.I. can't think independently and reason, but it can mimic arbitrator performance if trained by humans about the restrictions imposed on arbitrators. The aggregate of all these limitations leads to the supposition (to be confirmed) of a "perfect match," meaning that arbitration is a superior platform for integrating A.I. into the dispute resolution process.

A. The limitations imposed on arbitrators and A.I. make arbitration an ideal candidate for a computer-driven process.

These limitations on an arbitrator and on A.I. come into play when issues involving existing law require consideration. An established rule requires that unless the parties provide otherwise, the arbitrator can only apply the law as it actually exists and cannot add conditions because the arbitrator believes given circumstances appear to be slightly out of line with the law. In addition, the arbitrator isn't allowed to modify the law by considering evolving needs created by societal changes and pressures. Here are two examples of situations where these constraints are brought into sharp focus.

Suppose a buyer and seller contract for the purchase and sale of tires to be shipped to the buyer in Newark, New Jersey from Houston, Texas. The seller, in violation of the Jones Act, contracts to use a ship that isn't registered in the United States and is manned by citizens of the Philippines. The Coast Guard seizes the ship and holds it for ten days. The contract between the parties calls for arbitration "of any and all disputes" arising from their agreement to buy and sell tires. Query: Are the consequences of the delay caused by the seizure of the ship and therefore the substance of

[17]David Ha & Jurgen Schmidhuber, *World Models*, https://arxiv.org/abs/1803.10122 (2018).

a dispute within the meaning of the arbitration clause, *i.e.*, is this dispute "arbitrable"? And who decides the issue, an arbitrator or the court?

In *AT&T Technologies v. Communication Workers of America, et al.*.[18] the United States Supreme Court declared: "Unless the parties clearly and unmistakably provide otherwise, the question of whether the parties agreed to arbitrate (a substantive issue) is to be decided by the court, and not the arbitrator."[19] But the Court stopped short of addressing the question of "who" shall decide if there is clear and unmistakable evidence about the intent of the parties, and what standard must be used when making that determination. In *First Options of Chicago v. Kaplan*,[20] the U.S. Supreme Court established that the answer to this "who" question was the same: it depends on whether or not there is clear and unmistakable evidence that the parties want the arbitrator and not the court to decide.

Before the *First Options* case, an arbitrator had to defer to a court for an answer the question of whether clear and unmistakable evidence existed, and to do otherwise would have resulted in a declaration that the arbitrator had exceeded his or her authority. This would be so even if the arbitrator believed clear and unmistakable evidence actually existed. Similarly, a computer would have had to defer to a court for the answer. But the day after *First Options* was handed down, the arbitrator could determine the presence of clear and unmistakable evidence and therefore so could a computer.

First Options doesn't address the role that societal pressures play in the evolution of jurisprudence. Consider *Brown v. Board of Education,* 374 U.S. 483 (1954). This groundbreaking case and the abolishment of the "separate but equal" standard came about because the Supreme Court was willing to consider societal pressures and changing norms. In May of 1954, on the day prior to the handing down of the decision an arbitrator would have had to apply the "separate but equal" standard even if the arbitrator

[18] 475 U.S. 643 (1986).

[19] *Id.* at 656 (bracketed words added). What constitutes "Clear and unmistakable" evidence of what the parties want is yet another issue, one that has been handled on a case by case basis.

[20] 514 U.S. 938, 943 (1995).

believed that standard offended the 14th Amendment of the U.S. Constitution. Similarly, if charged to report on the law of desegregation as of the day before the *Brown* decision came down, the computer could do no more than recommend applying the "separate but equal" standard. If asked to review an arbitrator's ruling, issued the day before *Brown*, declaring the separate but equal standard unconstitutional, the computer would have to deem the ruling defective and an example of an arbitrator who has exceeded authority.

In the discussion that follows, we focus on two types of cases likely to be presented to a decision-making computer. *First Options* and *Brown* are both examples of a law case: one where the facts are not in dispute. The second is a fact case: one where the arbitrator and therefore the computer is called upon to determine the facts, determine the credibility and authenticity of written evidence and witness testimony, and apply applicable law. Here is the example:

> Hadley had his tailor hand-craft a dress shirt to be worn the day Hadley was scheduled for an audience with Queen Elizabeth. The tailor used only the finest quality fabrics and the shirt fit Hadley perfectly. It cost $450 plus sales tax at 7%. Hadley tried wearing it a week before his audience and inadvertently spilled mustard on the right sleeve. He took the shirt to his local dry cleaner and asked to have the stain removed. The dry cleaner accepted the shirt and issued a receipt containing an arbitration clause requiring arbitration before the Technically Savvy Arbitration Association, New York, New York using the Association's Commercial Rules then in effect. (Those rules are identical to the Commercial Rules of the American Arbitration Association (the "AAA").) When the shirt was returned, the left sleeve was badly burned. Hadley determined that the shirt was a total loss.
>
> The case manager has advised that it would be less expensive for the parties to submit the dispute to a specially equipped

arbitration model[21] containing an extensive, clean, and structured data set with over 100,000 dry cleaner burned shirt cases resolved by small claims courts throughout the New York Metro area and by arbitrators. The computer also has a complete library of New York law, including all reported judicial decisions of every kind, all New York statutes and governmental administrative rules and regulations of all kinds, and an extensive collection of secondary legal materials including digests and journals, form books, and other materials relied on by judges, arbitrators, and the arbitration administrators doing business in the New York Metro area. The case manager also advises that the model has been outfitted with extraordinary algorithms designed by the most sophisticated and respected computer scientists in the world, algorithms tested in simulated and real-world environments that have received international and national recognition and awards.

Hadley submits an affidavit swearing as truthful the shirt was brand new, never worn until the day he put it on and spilled mustard on the right sleeve. He states emphatically the left sleeve was in perfect condition when he left the shirt with the dry cleaner. The dry cleaner submits an affidavit claiming the shirt wasn't burned by any of the machinery in his store. He further swears that the pressing machine couldn't burn a shirt because of a fail-safe mechanism designed to shut the machine down if the temperature of the iron exceeds a safe level. Attached to his affidavit is a copy of the manual for the pressing machine and a statement from a repair man that the machine was in proper working order the day of the incident.

[21]For purposes of this paper, "model" is sometimes used to describe an algorithm and at other times an assemblage of two or more algorithms working in tandem.

B. A brief look at the technology involved with A.I.

Few professionals engaged in arbitration are also computer scientists, so getting the tangle of weeds that computer scientists routinely trouble over is not necessary for our purposes here. This discussion is focused on the practical, not the genius, involved in making A.I. a valuable tool.

For our purposes, A.I. is defined as the creation and use of basic algorithms designed to allow a computer to use training, data provided by the architect, to search through yet unseen sets of data looking for patterns and trends to answer a query with an almost certain probability. Overseeing and managing a model's activities are "predictive" algorithms, i.e. instructions telling the computer the steps needed for both dependent and independent performance.[22]

The science behind the creation of predictive algorithms is called "machine learning." There are three types:

1. Supervised learning,
2. Unsupervised learning, and
3. Reinforcement learning.

Deep learning is a subcategory of Unsupervised learning and today is the area receiving the most attention from the computer science community.

[22]The Association for Computing Machinery U.S. Public Policy Council defines an algorithm:

> An algorithm is a self-contained step-by-step set of operations that computers and other smart' devices carry out to perform calculation, data processing, and automated reasoning tasks. Increasingly, algorithms implement institutional decision-making based on analytics, which involves the discovery, interpretation, and communication of meaningful patterns in data. Especially valuable in areas rich with recorded information, analytics relies on the simultaneous application of statistics, computer programming, and operations research to quantify performance.

The Association identifies seven principles for transparency and accountability: awareness, access and redress, accountability, explanation. data provenance. Auditability, validation, and testing. *See, generally*, Ass'n for Computing Machinery, Statement on Algorithmic Transparency and Accountability (Jan. 12, 2017), https://www.acm.org/binaries/content/assets/public-policy/2017_usacm_statement_algorithms.pdf.

Deep learning involves the use of layered neural networks designed to mimic how human brains learn, by strengthening neurological pathways with repetition. Unlike Supervised learning that requires human control over data input, training, and the design of algorithms, Deep learning:

- Allows the computer to identify both apparent and hidden patterns embedded in very large sets of either labeled or unlabeled data, and
- Permits the machine to create algorithms necessary to make predictions.[23]

These models use pattern-recognition probabilistic methods.[24] To some this may sound like magic, but it actually involves credible and knowable mathematical methods to do this using a computer with suitable power to execute a huge number of calculations. The hope is that in short order new Deep learning methods will evolve allowing a computer to operate on its own, solving problems without any human support. Good data is at the core of any A.I. algorithm, and the quality of the algorithm is directly correlated to the quality of the data set used. The main function of A.I. is the unlocking of information and knowledge embedded in data. A.I. data only becomes useful if it's cleansed, well-labeled, annotated, and properly prepared for relevant input and analysis. This requires a significant investment in resources and time. Once data is prepared, it is ready for processing. See Figure 1 for further detail on the data lifecycle.

Computers can now receive, absorb, catalogue, and store data from many sources and in many forms, and can even receive and process data real-time without first having to store it. Advances in algorithm design now allow computers to:

- Interpret and understand human speech,
- Accept and read the written word,

[23]Harry Surden, *Machine Learning and the Law*, 89 Wash. L. Rev. 87, 94 (2014).

[24]Maxi Scherer, *Artificial Intelligence and Legal Decision-Making: The Wide Open? Study on the Example of International Arbitration*, SSRN.Com/abstract=3392669, at 8 and 36 J. Int'l Arb. 539 (2019).

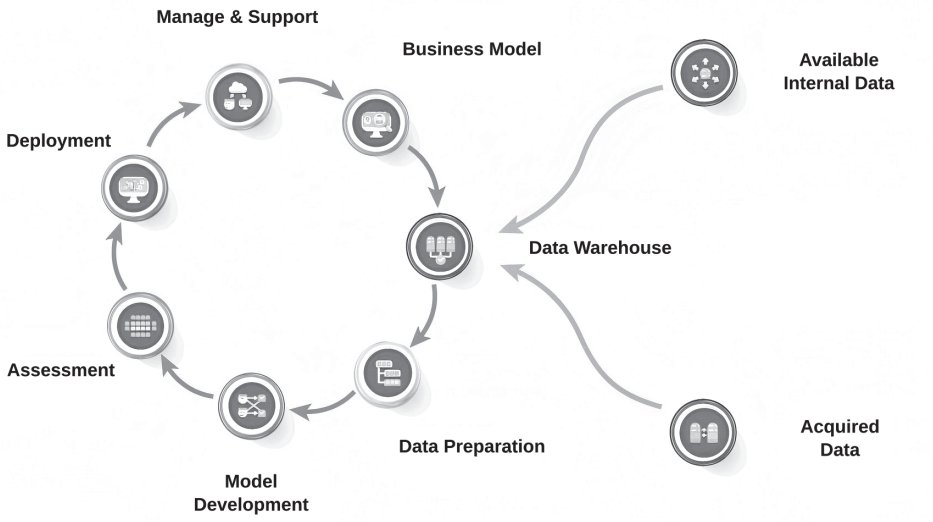

Figure 1. Cross-industry standard process for data mining

- Translate text to speech,
- Receive and interpret visual information, and
- Permit computers to independently monitor and update data from sources on the internet and elsewhere.[25]

With Supervised learning, data is labeled by a human and algorithms provide the required processing instructions. The techniques used have interesting names like "decisions trees" and "random forests," code words for finding a pathway leading to any one of the many possible end points. For example, an algorithm can be developed and trained to identify what a judge looks like using a large labeled data set composed of photographs of individuals exhibiting known features unique to members of the American judiciary: their robes, a gavel in hand, and perhaps gray hair and glasses. Using this data, the computer can be asked to determine whether a previously unseen picture of a person is a judge. With Supervised learning, a series of step-by-step instructions can require the computer to compare points in the picture one at a time, with labeled training data to support the conclusion of the presence of an element of a feature. If the comparison falls short, the computer is instructed to abandon the effort and a new search is ordered. If the comparison is successful, the computer is instructed to store the result and proceed to an analysis of the next point in the picture until sufficient information supports a final conclusion that the feature is present. With Deep learning, a very large unlabeled data set is used. The machine "learns" from patterns that are identified, stored, and then used to reach a conclusion the picture is or is not a judge, with a degree of certainty as close to 100% as possible.[26] No matter the system for learning, the end result is

[25]Amazon's recently launched the Alexa Prize competition (https://developer.amazon.com/alexaprize) requires competitors to create algorithms allowing the computer to have freeform conversations and be able to reference and converse about movies and recent news, all sourced from "live" data. The top performing teams achieved a maximum conversation lasting just over 8 minutes before the machine could no longer continue a "quality" conversation. *See* Jan Pichi et al., *Alquist: The Alexa Prize Socialbot*, https://arxiv.org/abs/1804.06705.

[26]While training an algorithm, successful determinations are stored in the computer's memory along with the information that led to each success.

a model that, when shown pictures not in the training data, is capable of answering the inquiry as to whether or not a picture is of a judge. Unfortunately, this model can't tell you if it's a picture of an arbitrator since it was only trained on pictures of a judge. It's important to note the overwhelming majority of A.I. tools are designed for a specific purpose and are not models worthy of broad application.

Machine learning techniques have evolved to a point where it is now possible to teach a computer a great deal about litigation and arbitration. However, some have raised the objection that data sets may contain embedded biases capable of undermining the ability of A.I. to be objective.[27] This is a justified concern, and we will discuss it later in this article.

For our purposes, assume any data set composed of arbitration related case law materials can be separated into three categories:

1. Cases involving a question of law,
2. Cases involving issues of fact, and
3. Cases involving both questions of law and issues of fact.

We will use *First Options of Chicago v. Kaplan* and *Brown v. Board of Education* as examples of cases involving a question of law. The computer's task is to understand the law involved. The case involving Hadley's shirt and the dry cleaner is one in which there are central questions of fact and credibility. As a practical matter, all cases "arbitrated" by a trained computer will involve questions of both fact and law.

Let's now look at some concerns with and questions about the concept of machine-driven dispute resolution and final arbitral awards.

---

[27]Caryn Devins et al., *The Law and Big Data*, 27 Cornell J.L. & Pub. Pol'y 357 (2017).

II. Question #1. How different is arbitration from courthouse litigation and why are the differences important?

Arbitration is *not* a substitute for litigation.[28] The only substantive similarity is a ruling on the merits. In litigation technical rules for pleading, discovery, and evidence are the norm. Not so in arbitration. In litigation, a final finding leads to a judgment and, in arbitration, an award. Judgments can be appealed as of right. Arbitration awards cannot be appealed with one exception: if the parties agree to submit the award to yet another arbitrator for review on the merits (this is very rare), whereas the finality of a judgment is not established until the right to an appeal is exercised or allowed to expire.

The most significant differences between arbitration and litigation involve reach and scope of authority. In the courthouse judges have broad discretionary powers. A judge's authority is found in the law they are sworn to apply and uphold. Not so for an arbitrator. The baseline for determining what an arbitrator can and cannot do is found in (a) the arbitration clause, (b) the limits imposed by statute, and (c) judicial interpretations of both the arbitration clause and controlling statutes. The arbitrator's first obligation is to the requirements specified by the parties, as long as their requirements do not offend public policy or are outright illegal. Judge Posner, quoted at the beginning of this article, reminds us that with the parties in full control, only trial by battle or ordeal is unacceptable.

Unlike in litigation, in arbitration the parties control several key issues by including or excluding any number of items in the arbitration clause. For example, they are free to provide rules for the selection of an arbitrator(s), rights to discovery, the law they want applied, the application of the rules of evidence, the timetable for the hearing, and the rules for the hearing.[29] They can even use a form of shorthand to provide for most of the items

[28] Judge Edward Weinfeld famously observed: "An arbitration tribunal is not a court of record; its rules of evidence and procedures differ from those of courts of record; its fact-finding process is not equivalent to judicial fact finding." *Williamson, P.A. v. John D. Quinn Constr. Corp.*, 537 F. Supp. 613, 616 (S.D.N.Y. 1982).

[29] For a comprehensive list, see American Arbitration Association, Commercial Arbitration Rules and Mediation Procedures, Preliminary Hearing Procedures P-2 (2013).

on this list. In the Hadley scenario, the clause is so short as to appear to be incomplete: "arbitration before the Technically Savvy Arbitration Association, New York, New York using the Association's Commercial Rules then in effect." In reality, this clause is complete because the provision incorporates by reference the commercial rules. And those rules provide that the parties can modify or eliminate any rule.[30] In Hadley's case, the applicable rules haven't been amended by the parties.

Significantly, there is no mention any place in the AAA Commercial Rules of the need for the arbitrator to apply law, meaning that unless the parties say otherwise, the arbitrator may do that which is thought to be reasonable. Similarly, absent a contrary direction from the parties, AAA Commercial Rule 34(a), "Evidence" provides "Conformity to legal rules of evidence shall not be necessary."

The ability of parties to exercise control over the arbitration process has implications for the application of A.I. Consider a situation in which the parties have required application of the laws of a given state. Selecting data sets containing those laws along with judicial decisions interpreting them is, in theory, feasible and simple. But if the parties fail to demand application of any specific law, the human arbitrator and therefore a mirroring computer would be charged to do what is thought by humans to be "reasonable." Selecting appropriate sets of data tailored to this situation may prove a daunting task given the unbounded debate over what the term "reasonable" actually means.

All commercial contracts that include an arbitration clause are automatically subject to the Federal Arbitration Act ("FAA").[31] Every state has its own version of an arbitration act, and parties are free to choose between the FAA and a state's arbitration act, subject, however, to the rule of federal preemption.[32] The FAA provides for the vacating of an award if the arbitrator runs afoul of Section 10(a) and all state arbitration acts have similar,

[30] Rule R-1(a) provides, in part: "The parties, by written agreement, may vary the procedures set forth in these rules." American Arbitration Association, Commercial Arbitration Rules and Mediation Procedures (2013).

[31] 9 U.S.C. §§ 1-16.

[32] *AT&T Mobility v. Concepcion*, 563 U.S. 333 (2012).

although not necessarily identical, provisions. Most important for our discussion is FAA. Section 10(a)(4) allows for vacating "where the arbitrators exceeded their powers . . . ." Arbitrator authority, discretion, and powers are limited by the agreement of the parties. Exceed those limitations and the award is subject to being vacated.

What constitutes exceeding powers is not always clear, and there's a large body of jurisprudence, federal and state, analyzing and discussing this topic. Libraries of this case law are readily available for training a computer. If a question arises about applying New York law, as opposed to some other law or no law at all, the computer would have to make that determination based on a review of a New York case law database. But suppose a computer isn't capable of deciding because no judicial teaching exists (the case is one of first impression), or there is a conflict of opinion between courts in the jurisdiction where the dispute is being heard? Algorithms can be designed to instruct the computer to signal the need for human intervention if the computer determines circumstances exist requiring the computer to perform beyond its scope.[33]

[33] *Asarco, LLC v. United Steel, Paper, Forestry, Rubber, Mfg., Energy,* 910 F.3d 485, 491 (9th Cir. 2018) involved a collective bargaining agreement with a "no add provision" restricting the ability of the arbitrator to "add to, detract from or alter in any way the provisions" of the agreement. After a hearing, the arbitrator found there was a mutual mistake concerning an amendment resulting in a dispute over the eligibility of certain persons to benefit from the agreement. Citing the mutual mistake, the arbitrator reformed the agreement to provide for the inclusion of the persons who might have otherwise been excluded. The court found that, because the arbitrator based his actions on an interpretation of contract law, the court lacked the ability to overrule what it found was an honest interpretation of contract law.

This case presents the question: What would a computer do if called upon to act as the arbitrator? The complexity of the arbitrator and the court's analysis suggests the computer may well be in over its head. Thus the need for an algorithm designed to have the computer signal the need for human intervention.

FAA Section 10(a)[34] creates a unique condition not found in courthouse litigation. If the human arbitrator's authority and powers are restricted, so must be the authority and powers of any computer mimicking the human arbitrator. Ensuring the faithful emulation of a human arbitrator is the task of the designer of an algorithm. The ability of the designer to confirm the presence of an appropriate algorithm should serve to mollify fears that an algorithmic-driven computer might be dispensing what the computer, or its human masterminds, believe to be their version of justice.[35]

III. Question #2. Is a computer really just a black box? Exploring and resolving issues of embedded bias.

On numerous occasions courts have held that parties in an arbitration aren't entitled to a perfect hearing. They are, however, entitled to a fair hearing.[36] What constitutes a fair hearing isn't always an easy measure, but at the very least, the arbitrator must operate in the open where the parties can

[34]Section 10(a) provides: (a) In any of the following cases the United States court in and for the district wherein the award was made may make an order vacating the award upon the application of any party to the arbitration:

> (1) where the award was procured by corruption, fraud, or undue means;
>
> (2) where there was evident partiality or corruption in the arbitrators, or either of them;
>
> (3) where the arbitrators were guilty of misconduct in refusing to postpone the hearing, upon sufficient cause shown, or in refusing to hear evidence pertinent and material to the controversy; or of any other misbehavior by which the rights of any party have been prejudiced; or
>
> (4) where the arbitrators exceeded their powers, or so imperfectly executed them that a mutual, final, and definite award upon the subject matter submitted was not made.

[35]State arbitration statutes all provide conditions for vacating an award. Most follow the federal statute but there are exceptions. For example, the New York C.P.L.R. allows vacating upon a showing that the arbitrator failed to follow the procedure set forth in Article 75. *See* N.Y. C.P.L.R. 7511(b)(1)(iv). In addition, a finding by an arbitrator that the claim was barred by a statute of limitations is grounds to vacate. *See* N.Y. C.P.L.R. 7511(b)(2)(iv).

[36]*CM South East Texas Houston, LLC v. CareMinders Home Care, Inc.*, 662 F App'x 701, 705 (11th Cir 2016), *Employers Ins. Of Wausau v. National Union Fire Ins. Co.*, 933 F.2d 1481, 1491 (9th Cir 1991), *Hoffman v. Cargill*, 988 F. Supp. 465, 474 (N.D. Iowa 1997).

observe demeanor and professionalism. No arbitrator can operate behind a curtain.[37] Transparency, meaning overt behavior consistent with a commitment to neutrality and freedom from prejudice and bias, presents unique complications for the use of A.I. in an arbitration. While a computer can be programmed to resolve a problem, for the moment, algorithms directing the computer to explain to the user why it is doing what it is doing are still a work in progress.[38] So, for now there is a perception of computers being a "black box" that may be doing the bidding of someone other than the user. If the computer cannot account for its actions, how can a user know that data sets and algorithms are not tainted by undisclosed, undetectable and/or unintentional biases?[39]

To put this into a proper context, first consider whether human arbitrators are likely to be biased. There's no shortage of cognitive psychological

[37]Andrea Roth, *Trial by Machine*, 104 Geo. L.J. 1245 (2016). See, in particular, the "man behind the curtain" discussion at pp. 1277-80.

[38]Ryan Calo, *Artificial Intelligence Policy: A Primer and Roadmap*, 51 U.C. Davis L. Rev. 399, 414 (2017). "No one knows why the system selects a document: once the system is trained, no script can be provided to a human sorter to imitate the system's selection of documents. That is, there is no way to accurately summarize the criteria used. Nevertheless, parties rely on predictive coding in very high-stakes litigation. It is treated as reliable." Curtis Karnow, *The Opinion of Machines*, 19 Colum. Sci. & Tech. L. Rev. 136, 142 (2017). "In general, deep learning models (that is, deep artificial neural networks) are often criticized as of being "black-box" models, whose answers, despite being remarkably accurate, are hard to interpret. There is a major need, in the field of AI, to build explainable models, i.e., models capable of motivating their choices, that is models whose decision processes can be interpreted by a human. The direction in which the field is moving is that of integrating so-called sub-symbolic (or connects) approaches, such as artificial neural networks, with so-called symbolic methodologies, which are built on logic. The former are capable of efficiently and effectively dealing with uncertainty in data and can easily exploit very large data collections, but lack in interpretability. The latter, on the other hand, are designed to deal with knowledge representation and reasoning, and thus show a high expressivity, a high interpretability, but cannot easily handle noisy information and scale to big data. There is a strong belief within the AI community that the combination of such diverse approaches is a necessary step to fill the performance gap in tasks related to reasoning." Przemyslaw Palka & Marco Lippi, *Big Data Analytics, Online Terms of Service and Privacy Policies*, https://ssrn.com/abstract=3347364, at 21-22 (internal citations omitted); *see also* Lisa Getoor & Ben Taskar, *Statistical Relational Learning* (2007); Artur d'Avila Garcez et al., *Neural-Symbolic Learning and Reasoning: Contributions and Challenges*, 2015 AAAI Spring Symposium Series.

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

studies pointing to "yes." Through observation, parties can detect some biases. An arbitrator who shows a preference or distaste for someone based on race, religion, gender, or some other unique factor should be easy to spot. For example, a showing that 30 out of 30 awards favor a white complainant over a black respondent does justify inquiring if the arbitrator harbors racial animus. But sometimes it's not so clear. For example, a pattern of awards disproportionately favoring banks over consumers seems to suggest a bias. But there may be other explanations. Perhaps the reason has to do with the number of respondent consumer defaults leaving the arbitrator no choice but to rule in the bank's favor.[40] The takeaway is that while a pattern of behavior may point to a known bias, there is still a need to use caution and do further analysis.

Cognitive unconscious biases are a different matter. Empirical studies have shown that arbitrators, judges, and juries bring to their roles hidden biases that often they themselves are unaware of. These biases, some call them blinders,[41] result from the human tendency to use heuristics—mental shortcuts—when making decisions.[42] Most individuals are totally unaware since these biases are embedded in the unconscious and are often further supported by cognitive dissonance. Anyone asking an arbitrator about an unconscious bias is not likely to get meaningful information because the arbitrator, unaware of the unconscious process, will in good faith deny falling victim to bias. For instance, a "coherence" heuristic was identified in a study involving judges who were asked to estimate their reversal rate relative to their peers. The study showed that judges are inclined to believe their own rulings are correct at a rate exceeding that of their peers. In the study, judges had to estimate their reversal rates by an appellate court. Fifty-six percent rated themselves in the lowest reversal rate group and 31% rated themselves in the next lowest reversal rate group. These results suggest

[40] Paul Bennett Marrow, *Counsel, Beware*, 81 NYSBA J. 36 (2009).

[41] Chris Guthrie, *Misjudging,* 7 Nev. L. Rev. 420 (2007).

[42] *See* Daniel Kahneman & Amos Tversky, *Subjective Probability: A Judgment of Representativeness*, 3 Cognitive Psychol. 430 (1972) [hereinafter "*Subjective Probability*"]; Daniel Kahneman & Amos Tversky, *Availability: A Heuristic for Judging Frequency and Probability*, 5 Cognitive Psychol. 207 (1973); Daniel Kahneman & Amos Tversky, *Judgment Under Uncertainty: Heuristics and Biases*, 185 Science 1124 (1974).

that 81% of the judges' thought that at least half of their peers had higher reversal rate records than they had.[43] But if you were to ask any of the participants if he or she had inflated their perception of judicial acumen, the general answer, given in good faith, would be "of course not." Another example involves studies showing how the "anchoring effect" heuristic impacts the ability to estimate an unknown quantity. If a participating judge attempting to settle a matter is given a number, even if this number is random, the participating judge will likely anchor to the provided number, making any estimate highly unreliable.[44] If asked if there is any factor influencing their determination as to what constitutes a reasonable number for settlement purposes, the most likely answer might be "no."

Factors unknown to a judge, even factors that seem totally irrelevant to the dispute at hand, can influence decisions. For example, a famous study involving Israeli judges hearing applications for parole revealed a direct connection between the likelihood of securing parole and the timing of a judge's lunch break. While applications for parole were infrequently granted, it turned out chances greatly improved at the very beginning of the working day and again after a judge's lunch break.[45]

Influencing and/or defining the algorithm design process is fertile ground for the insidious involvement of unconscious biases. The same is true for training data.[46] An algorithm needs a structure and the determination of the structure can easily involve a subjective criterion not apparent to the author or those using the output. To state the obvious, how an algorithm processes data dictates the objectivity and value of the output.[47] For exam-

[43] *Subjective Probability*, *supra* note 42, at 436-37

[44] Chris Guthrie et al., *Inside the Judicial Mind*, 86 Cornell L. Rev. 777, 792-94 (2001) ("The potentially pernicious effects of anchoring also suggest a source of error in both the civil and criminal justice systems. In civil cases, the influence on judges of misleading anchors, such as litigants' for damage awards, can produce biased damage awards.").

[45] Shai Danziger et al., *Extraneous Factors in Judicial Decisions*, 108 Proc. Nat'l Acad. Sci. 6889 (2011), https://www.pnas.org/content/108/17/6889.

[46] What is Garbage In, Garbage Out (GIGO)?, https://www.techopedia.com/definition/3801/garbage-in-garbage-out-gigo.

[47] Scherer, *supra* note 24, at 21. Hidden bias is not the only concern. Random errors caused by mistyping and mistranslation are examples of other problems. Andrea Roth, *supra* note 37, at 1275.

ple, consider an algorithm designed for learning from data being contaminated because it allows for the inclusion or exclusion of certain parameters. The scope of this problem, however, is somewhat contained by the ability of the architect to physically reexamine the design of any algorithm. More insidious is the potential for the failure or inability to properly manage and screen input data.

Training data originates from two sources. It can be selected by third parties and/or by the architect of the algorithm. A major risk is training data contaminated with unconscious biases of the people involved in the selection process. One observer noted it's easier to detect error/bias in algorithm design than detect and correct them in training data sets. Since learning models also retrain and reinforce using prior results obtained with tainted data, the biases are likely to become self-perpetuating, further detracting from the value of the model.[48] As a result, there is a real risk over time that undetected data set errors and biases can become so deeply embedded as to take on meanings of their own and eventually change the algorithm without human detection. The experience of Staples provides an early example of our blindness to the outcomes of algorithm and data selection designs containing imbedded, albeit unintentional, bias. Staples deployed an algorithm that unwittingly discriminated against certain consumers based on social and economic status. The root of the problem was traced to the training data. Staples identified an unintended bias allowing the offering of reduced prices to buyers in more affluent neighborhoods of means when the purpose of the algorithm was to offer the lower prices to a less privileged population. The investigation discovered that the training data contained an assumption that those living closer to a brick and mortar location of a competitor such as OfficeMax would be less affluent and therefore more price sensitive than those living further away. It turned out that less affluent buyers actually lived further from the competition and yet were shown higher prices than the more affluent living nearby the competition.

[48]Scherer, *supra* note 24, at 559-61.

The final conclusion was the algorithm was relying on skewed training data that contained an undetected bias favoring the affluent.[49]

Consider the example given earlier of training a computer to recognize a judge. If the photos shown to the computer include only elderly men with gray hair, glasses, and black robes, the computer, relying on the training data, will fail to identify pictures of young men and women and older women, no matter the color of their hair or their garb, as judges.[50]

Data may be biased simply because of outside and/or societal pressures that have helped to shape the data. For example, if the data contains information that is correct and yet unbalanced, any prediction made using that data will be unbalanced. Assume data sampled shows a ratio of four (4) apples to every one (1) orange, and an overall likely rotting rate of 10%. It follows that the outcome will reveal far more apples have rotted than oranges. Now consider a sampling of convicted criminals 85% who are black and 15% who are white used in a study to determine the likelihood of a convicted criminal committing another crime once released from prison. It's a certainty that the resulting statistics will show more recidivism among blacks in the sample than whites. As one observer put it: "The outcome is biased because reality is biased."[51]

Designers have a humbling responsibility that must be taken very seriously.[52]

What is being done to address these concerns? The simple answer is "a lot," but the solutions available to date are only a beginning. Research has led to a number of possibilities.[53] One suggested solution is asking the developer of the data set and/or algorithm for an opportunity to audit the inner workings before anything is deployed. While an examination of the

[49]Jennifer Valentino-DeVries, *Websites Vary Prices, Deals Based on Users' Information*, Wall St. J. (Dec. 24, 2012).

[50]*See* Amanda Levendowski, *How Copyright Law Can Fix Artificial Intelligence's Implicit Bias Problem*, 93 Wash. L. Rev. 579, 591-92 (2018) (discussing a computer being trained to identify a cat).

[51]Hannah Fry, *Hello World: Being Human in the Age of Algorithms*, 68 (2018).

[52]For an in-depth discussion, see Harry Surden, *The Ethics of Artificial Intelligence in Law: Basic Questions*, (Aug. 26, 2019), www.ssrn.com/abstract=3441303.

[53]Levendowski, *supra* note 50, at 586 nn. 26-27 (2018).

inner workings of an algorithm or the criteria for data set selection should expose errors and biases of all types, there is no guarantee.[54] Moreover, as a practical matter, there is a likelihood of resistance if the author is concerned about the trade secret value of the work. (Some algorithms are open source.)

Another approach is testing using queries designed to show the existence of an embedded bias. If a bias is detected, asking the author to account for it would allow for an explanation and even a solution and is unlikely to be threatening so long as there is no need for the disclosure of sensitive information. However, care must be taken to ensure that any testing techniques are broad enough in scope to be able to detect a broad range of possible biases.

Data sets present yet another challenge. Litigators frequently search on the name of a judge for any insights into how the judge is likely to rule on an issue. Given the nature of arbitration and the emphasis on confidentiality, the reality is that few arbitrators file awards that are available for public review. Some issue unreasoned awards, meaning that an award does not include an explanation of the reasoning behind the decision. And most arbitrators issue awards involving an assortment of topics. For these reasons, it is very difficult to test for a pattern suggesting an individual's bias.

Beyond testing, much effort is being given to what computer scientists call "explainability." It is commonplace for parties to demand a human arbitrator produce a reasoned decision, one that reveals the thinking of the arbitrator and explains the details about the law applied and the evidence found credible. Demanding an algorithm capable of providing this level transparency seems only reasonable. Is the science involved in A.I. capable of directing a computer to explain any award it issues? Programming models to provide an explanation of the parameters used for deciding is a nascent area of research, but one that has shown early promise.[55] The tradeoff has been between "explainability" and effectiveness. While explainable

[54]Robert Brauneis & Ellen Goodman, *Algorithmic Transparency for the Smart City*, 20 Yale J.L. & Tech. 103, 117-20 (2018).

[55]Wojciech Samek et al., *Explainable Artificial Intelligence: Understanding, Visualizing, and Interpreting Deep Learning Models* (Aug. 28, 2017), https://arxiv.org/abs/1708.08296.

A.I. exists (commonly referred to as XAI), the level of "explainability" is inversely proportionate to the complexity of the problem and the deployed model used to solve it. Simple statistical models, like those currently used to determine insurance premiums, credit card rates, or loan approvals are typically based on a decision tree analysis. But for more complex systems, like Deep learning, the research is only now beginning to show progress. As was noted earlier, Deep learning involves algorithms allowing a computer to evaluate data unsupervised. The decision process employed by the computer can be poorly defined and at times appear to the human being as based on nothing more than useless noise. In addition, because a neural network is multi-layered, tracing back the decision process is an undertaking that is highly complex, time consuming, and expensive.

A second concern is determining the magnitude of explainability within a given setting. The ethical and legal issues involved and the degree of explanation required differ for every application of A.I. While the concern has led to discussions among those designing algorithms that permit explanation, among those using them, and those who are considering how to regulate the need for explanations, there is still no agreement on how to set parameters flexible enough to embrace the full scope of a given A.I. application.[56]

A third area of concern involves expectations. Given the concern about the "black box," in an arbitration case, the ideal would be an audit function providing details about the computer's decision-making process indicating what factors, evidence, and law is being considered as well as the factors, evidence, and law disallowed. The ability of the science behind XAI to deliver this type of explanation is not yet fully developed. Most of the work in XAI has involved "simplified approximations of complex decision-making functions." These approximations appear to users to be more like scientific models than the "everyday" explanations the user community is looking for.[57]

---

[56] Adey Zegeye, Design/Ethical Implications of Explainable AI (XAI) (May 7, 2019), https://blogs.commons.georgetown.edu/cctp-607-spring2019/2019/05/07/design-ethical-implications-of-explainable-ai-xai/.

[57] Brent Mittelstadt et al., *Explaining Explanations in AI* (Nov 4, 2018), https://arxiv.org/abs/1811.01439.

Yet another area of concern is the philosophical implications of the differences between how humans and computers perceive and evaluate a dispute. Scherer[58] observes that humans have developed systems of law that are serviced through deductive reasoning, logic, and the application of known rules. This allows for some flexibility, as required by the circumstances of a specific dispute. A.I., on the other hand, is currently designed to resolve an issue using mechanistic formulas grounded in pattern identification and statistical probabilities derived from the learning data. No matter the training method or the model, A.I. can do little more than provide a predictive output as to a degree of mathematical certainty. Acceptance of this type of protocol could prove a hard sell since it would require users to knowingly abandon the ingrained notion that law is a function of deductive reasoning, logic, and application of known rules.[59]

On the other hand, the use of deductive reasoning and logic can allow the arbitrator a degree of flexibility that at times can entice an arbitrator to exceed authority, something § 10(a)(3) of the FAA prohibits. For example: The arbitrator is confronted with a case of first impression, i.e. a situation that no judge has ever considered and ruled on. An argument can be made that the arbitrator can only apply a known law and lacks authority to fashion new law to accommodate such a situation and that doing anything else amounts to dispensing his or her own brand of industrial justice. This risk is reduced substantially, if not eliminated altogether, using a computer

[58]"[S]ome computer models (such as expert models) are indeed rule-based, using causal logic and deductive reasoning, since they apply pre-established rules in the algorithm to the observable data. Other AI models, however, have different features. In particular, machine learning models, such as neural networks, often have no pre-defined rules. Deductive, causal reasoning is thus replaced by an inverse approach, because the machine learning program extracts the algorithm from the observable data. Rather than using logic, the AI model calculates probabilities, i.e. the likelihood for any given outcome. [¶] Applying such machine learning processes in the legal decision-making context therefore would mean accepting a departure from the above-mentioned understanding of judicial reasoning according to formalist theories. A decision based on those AI models would not be based on pre-determined legal rules, would not be the result of deductive logic, and would not follow the above-described legal syllogism. While this situation would be a cause for concern for legal formalists, it might be seen as vindicating others who have long criticized formalist theories." Scherer, *supra* note 24, at 567.

[59]For an in-depth discussion, see *id.* at 562-72.

trained with a proper database and equipped with algorithms designed to have the computer alert the designer if confronted by such a situation. In addition, the computer's conduct is devoid of human emotion and subjectivity. Over time these considerations may prove sufficient to allow arbitration by computer to become an acceptable alternative.[60] And finally, arbitration training data is anchored to the law applicable to arbitration. While the computer may be looking for patterns and applying mathematical formulas to obtain a probability, it is doing so using the same materials an arbitrator would when making a judgment leading to an award.

What is clear is the entire field involves many complexities,[61] and there is a need for much further research and work. While it is anticipated and expected that most applications of A.I. will eventually become explainable, we may not be able to answer whether or not the explanations themselves will be sufficient, auditable, and trustworthy.

Beyond testing and XAI, consider using a panel of three independently trained computers or two independently trained computers and one human being. Presumably the panel, however composed, would vote and the major-

[60] *Id.* at 557, worries that the reliance on data composed of previous decisions, there is a risk that computer results will be "conservative" and not be adapted to policy changes over time: "Velocity refers to the frequency of incoming data that needs to be processed. Big Data is often challenging because of the sheer amount and high frequency of the incoming data. In the legal context, such risk is very low. As already pointed out above, in terms of volume, the problem is likely to be of scarcity rather than abundance of data. Therefore, over time, decisions might not be frequent, and when they occur there might have been a change in policy so that the previous data is outdated. These policy changes can be radical and swift at times. To take an example from the international arbitration context, the decision of the Court of Justice of the European Union in Achmea has fundamentally changed the compatibility of investor-state arbitration with European law overnight. This raises the question how AI models which, by definition, are based on information extracted from previous data may deal with those policy changes. It is true that the essence of machine learning is the ability to improve the algorithm over time. Nevertheless, such improvement is always based on past data. Policy changes in case law necessarily require departures from past data, i.e. previous cases. For these reasons, AI models are likely to keep conservative' approaches that are in line with previous cases." Scherer's concern disregards the reality that an arbitrator is rarely, if ever, endowed with the authority needed to fashion a rule needed to address evolving policy considerations.

[61] Samek et al., *supra* note 55; Finale Doshi-Velez et al., *Accountability of AI Under the Law: The Role of Explanation*, Harvard Law School, Public Law & Legal Theory, Research Paper Series, #18-07 (Nov. 6, 2017), https://ssrn.com/abstract=3064761.

ity would prevail. The format involving two computers and a human being seems to defeat the purpose of turning decision making over to machines.[62]

Some have challenged the appropriateness of using a "big data" set. The concern is: (1) Big data isn't objective. (2) Big data doesn't consider the evolutionary nature of our law and legal system. (3) Big data risks failing to reflect the true nature of our legal system and instead will reflect a system all its own.[63] These concerns reflect a misunderstanding of the role big data plays in Machine learning and, in particular, unsupervised Deep learning. A well-trained Deep learning computer receives data from any number of diverse sources. The human element is absent from the selection and labeling of data thereby eliminating human source subjectivity. The objection that data might not consider the evolutionary nature of our law shows a misunderstanding of the role A.I. can play in arbitration. Unless allowed by the parties to an arbitration, an arbitrator can never consider changes in societal norms. Indeed, this constraint is one of the most compelling reasons why arbitration is a suitable candidate for A.I. driven programs.

Whether a human or a computer, there is a risk bias will play a role in the decision-making process. But that doesn't necessarily mean an award is unfair. What gives the computer the edge is the ability to uncover and remove bias using simulation techniques.

What might a data set tailored to Hadley's case look like? The design of an appropriate data set would entail a great deal of forethought to ensure its appropriateness. The peculiarities of the basic elements of the case, (let's call them generalized data points) should drive the criteria for the data sets needed to train a computer. The particular model would need to be trained to understand what a shirt like the one involved looks like, a description of the type of fabric(s) used, what mustard is, how it stains fabric, what a mustard stain looks like, what the cleaning process is, what can cause a

[62]For an example where this type of team approach has had positive results, see Katyanna Quach, *Don't try and beat AI, merge with it says chess champ Gary Kasparov*, The Register (May 10, 2018), https://www.theregister.co.uk/2018/05/10/heres_what_garry_kasparov_an_old_world_chess_champion_thinks_of_ai/.

[63]Caryn Devins et al., *The Law and Big Data*, 27 Cornell J.L. & Pub. Pol'y 357, 359-60 (2017).

burn on fabric during the cleaning process, what the tell-tale signs of a burn caused by a machine such as the one used by the dry cleaner are, what a burned shirt sleeve looks like, what a shirt that isn't burned looks like, what a dry cleaner is and how they operate, and what is the typical relationship between a dry cleaner and a customer. Photographic information if any, showing the shirt and its condition at any point in the timeline leading to the dispute would also be required. Other general data points might include information about the machine used to clean and press the garment along with the manual(s) describing any fail safe mechanism claimed to have been running the day the shirt was allegedly damaged and information about the proper process for removing a mustard stain. If available, a data set composed of cases involving damage claims against dry cleaners and verdicts together with a library of photographs showing similar burns might be considered. Other general data points to be considered might include information about either or both parties' prior lawsuit/arbitration activity. Addressing issues of credibility might require developing algorithms trained on data sets composed of examples of shirt owners and dry cleaners who are known to either be lying or telling the truth.

Dry Cleaner consumer disputes occur throughout the U.S. Each jurisdiction charged with resolving these disputes has its own rules and laws that often vary. Some jurisdictions may not have a population able to afford to purchase a shirt of the quality of the one Hadley bought, raising the question of whether or not it is appropriate to include a "proxy" for such shirts. If shirts aren't the measure, should other garments similarly priced be used as a proxy? While a $450 shirt may not be the norm, a $450 men's suit might be. The point becomes obvious; great care is needed when determining the scope and nature of the data.

IV. Question #3. How complete need the data set be before it can be used?

The more robust a data set the better A.I. and the computer will perform.[64] No data set ever is complete, there is always room for additional information.[65] Within the context of arbitration, having on hand all the relevant "book knowledge" available through services such as LEXIS and Westlaw and more is important. Access to most, if not all of this data is not a problem. However, finding "real" substantive information about the details of most arbitration proceedings is likely to prove to be difficult. Hearings are closed, and transcripts, if made, are confidential unless a party seeks to vacate.[66] FINRA makes its awards available, unredacted, online, but doesn't reveal the supporting materials such as the exhibits and memorandums of law submitted by the parties or motion practice materials. Redacted awards issued by AAA Labor and Employment arbitrators are available at Lexis and Westlaw. Bloomberg Law makes available numerous international awards. Westlaw has a library of awards submitted by insurance arbitrators and a library containing international awards. Most awards that are available contain the name of the arbitrator as well as the result reached. When a party seeks to vacate an award, what is provided to the court is publicly available. Another source are the annual filings by any facilitator referred to in this article conducting business in California, Maryland, Maine, and the District of Columbia. While these filings are not complete, they do reveal the names of arbitrators, the number of cases heard, and the awards made.

In the past few years, courts throughout the U.S. have opened their files to the public at sites on the internet. While procedures in the courthouse may differ, these files should contain information relating to petitions to

[64]Scherer, *supra* note 24, at 15 ("Any data-driven AI programs first and foremost require access to data. Machine learning models, which are based on probabilistic inferences, are data hungry: the larger the sample data, the more accurate the model's predictive value.").

[65]Elite Data Science, Overfitting in Machine Learning: What It Is and How to Prevent It (2019), https://elitedatascience.com/overfitting-in-machine-learning.

[66]Confirmation is the process whereby an award is converted into a judgment. Confirmation typically doesn't require the record of the hearing. An application to vacate does require a fairly comprehensive description to the court of the issues and facts formulating the basis for the demand. *See, e.g.*, 9 U.S.C. § 13.

confirm or vacate awards. Compiling a library with this information may prove expensive, but worthwhile.

As efforts are made to train computers using existing data sets, testing will no doubt reveal deficiencies needing to be addressed. Thinking through the details of any arbitration is arduous and time consuming for the parties and counsel. Training a computer with algorithms and data sets needed to address a broad range of conditions and situations is a far more complex task. Template style data sets describing in general terms the common factual elements of any number of disputes alone may not suffice. The training effort will probably require the development of algorithms designed to respond and provide structures addressing the peculiarities of any individual case.

Selling the idea of a machine-driven binding award may prove a daunting task. Along with transparency, the completeness of the data sets and the appropriateness of governing algorithms are likely to be two of the most controversial aspects of any program. Adding to the difficulties, the public will have to be educated about the use and meaning of any number of terms employing examples that are as non-technical, and yet as persuasive, as possible.

V. Question #4. What kind of cases would be best served by a computer acting as the arbitrator?

Today, the science involved with machine-driven final binding awards is at an embryonic stage. The types of disputes and legal issues being considered for arbitration are ever-growing and may prove to be without limit. Still, as a practical matter, not every dispute is a suitable candidate for A.I. There will always be a self-selecting process evidenced by the earlier suggestion that stakeholders in a "bet the company" case will more often than not reject the idea of machine-driven arbitration. But what about all the other disputes of great importance to the parties, even if the financial consequences aren't draconian? Will all or just some be candidates for arbitration using a computer? In all likelihood, most disputes will prove to be non-compatible. The more complex factually or at law a dispute becomes, the less likely the disputants and in particular counsel will accept a machine-driven process. Complex disputes usually require the absolute right to discovery in all al-

lowable forms. The same is true concerning the rules of evidence. These rights are not automatically available in an arbitration.

The paramount benefit attributed to arbitration is a desire to keep things as simple and streamlined as possible. The default position is dispensing with as many courthouse formalities as possible and bringing the case to a conclusion as quickly as possible. Given the preference for simplicity, the best candidates are likely small claims matters, defined as involving disputes (1) at law, excluding equitable matters of any kind, (2) with a dollar value of no more than a set amount, probably less than five figures, (3) involving simple factual and legal issues (4) and limited legal defenses, and (5) that can be easily classified as typical. For example, a suit on a promissory note is usually straightforward. The debtor either hasn't paid and has no excuse or hasn't paid and has an excuse that can be established with document evidence. Also included are claims involving limited property damage, failure to comply with a clearly defined obligation, breach of a contract claims involving the delivery of goods or services, and minor negligence matters. In Hadley's dispute, the underlying facts and law involve the basic elements of bailment and negligence. To prevail Hadley must, at a minimum, show:

- Ownership of the shirt, including proof of payment
- His care for the shirt prior to having it cleaned
- The value of the shirt prior to having it cleaned
- Proof that the shirt wasn't burned before it was handed over
- The actual condition of the shirt when handed over
- A description of the take-in procedures involved when the shirt was turned over

The dry cleaner, in order to defeat the claim, will have to show at a minimum:

- What the take-in procedures were when the shirt was presented

- What the custom and usage standard was at the time among dry cleaners in the community concerning take-in procedures
- What method was used to clean and press the shirt
- What type of equipment was used to clean and press the shirt
- The condition of the equipment on the day and time when the shirt was cleaned and pressed
- What fail safe mechanisms or procedures were in place to avoid damage to garments processed
- What the standard of care by dry cleaners was at the time of Hadley's transaction
- Identification of the individuals at the dry cleaner who were involved at the time Hadley's garment was taken in and subsequently processed

These elements are likely to be present in any dispute involving a claim of damage to property left with a dry cleaner.

Similarly, the duties imposed by law in this type of claim are elementary. The dry cleaner's duties are to inspect and take in the garment using procedures common to dry cleaners in the community, processing the shirt in a reasonable manner to avoid damage, and finally to ensure the return of the cleaned garment in substantially the same condition as when received. The standard of care is reasonableness under the circumstances.

In Hadley's case, the human arbitrator's first task would be to evaluate the information provided and ask questions designed to (a) ensure a complete picture of what actually happened to the shirt while in the possession of Hadley and then the dry cleaner, and (b) establish the credibility of witnesses and tangible evidence offered to establish or defeat the claim. An arbitrator might want to know:

- Before this claim was there a history of disputes between Hadley and this dry cleaner?
- Did either party have a history suggesting a penchant for litigation?

- Did anyone other than Hadley see the shirt and evaluate its condition before Hadley brought it to the dry cleaner?
- Why did a repairman examine the equipment on the day of the incident?
- Who, if anyone, other than the dry cleaner himself actually handled the shirt when it was cleaned and pressed?

To ensure the legal issues being decided and factual contradictions needing resolution meet the requirements for arbitration by computer the parties would have to submit a "package" of materials for review by the provider of the service. If determined to be insufficient or incomplete, the party involved would have to address the deficiency before the matter could be presented to the computer for determination.

Other types of disputes might also lend themselves. Insurance companies have created mechanisms for resolving disputes concerning distributing liability between carriers who have insured a risk. Many of these disputes are pro forma involving few complexities. These disputes, along with others of a similar nature arising in the business to business community, are candidates with the benefit being cost savings realized by the elimination or substantial reduction of the human factors in the dispute resolution process.

VI. Question #5. Will it be necessary to amend or modify the FAA or any of the state arbitration laws?

The short answer is "probably."

The FAA doesn't define the term "arbitrator." Nevada is the only state with an arbitration statute that defines the term.[67] Few courts have looked into the issue. The U.S. Supreme Court, as far back as 1868, opined "An arbitrator is defined as a private extraordinary judge chosen by the parties who have a matter in dispute, invested with power to decide the same."[68]

[67] Nev. Rev. Stat. § 38.209 ("'Arbitrator' means an individual appointed to render an award, alone or with others, in a controversy that is subject to an agreement to arbitrate").

[68] *Gordon v. United States*, 74 U.S. 188, 194 (quoting *Bouvier's Law Dictionary*, title "Arbitrator.").

While the Court in 1868 wasn't aware of computer technology and algorithms, the reference to a "private extraordinary judge" is quite broad and arguably could include new technologies. The Supreme Court of Wisconsin accepted the definition offered by *Webster's Third New International Dictionary (1967):* "one with absolute power of deciding disputes so as to bind the disputants"[69] (Can "one" be interpreted to include a computer?) Most other state courts speaking about arbitrators as actors refer to them as being a "person."[70] Perhaps most important though, FAA Section 10 makes no reference to an arbitrator being a person. The described missteps sufficient to warrant vacating do not appear to provide a basis for concluding Section 10 would have no application just because the arbitrator is a computer. However, there is good reason to wonder if a misstep by a person who has trained the computer would come within the scope of Section 10. Given this question alone, it appears some legislative action might be needed to insure recognition of a computer as an arbitrator within the meaning of the Act.

None of the rules of the major arbitration administrators[71] address the issue of who can serve as an arbitrator. FINRA rules define both a public and a private arbitrator as "a person."[72] The Business-to-Business rules of the National Arbitration Forum define an arbitrator as:

---

[69] *Grays Harbor County v. Williamson*, 96 Wn. 2d 147, 156 (1981).

[70] *State ex rel. Cushion v. City of Massillon*, 2011 Ohio App. LEXIS 3922; *Konz v. Morgan Stanley Smith Barney,* 2018 U.S. Dist. LEXIS 180069*; Bolick v. Merrill Lynch, Pierce, Fenner & Smith, Inc,* 2006 U.S. Dist. LEXIS 330; *Nieves v. Travelers Cas. Ins. Co. of Am.*, 2015 U.S. Dist. LEXIS 95774; *Society of Am. Foresters v. Renewable Natural Resources Found.*, 114 Md. App. 224 (1997); *State ex rel. Cushion v. City of Massillon;* 2011 Ohio App. LEXIS 2457*; Hino Motors Mfg. United States v. Naftaly*, 2011 Mich. App. LEXIS 1562.

[71] For purposes of this paper, these include the AAA, JAMS, CPR, NAF, and FINRA.

[72] *See, e.g.*, FINRA Arbitration Rules 12100 (r) and (y). JAMS Comprehensive Arbitration Rules and Procedures, Rule 7 (a) states: "In these Rules, the term 'Arbitrator' shall mean, as the context requires, the Arbitrator or the panel of Arbitrators in a tripartite Arbitration."

> An individual selected in accord with the Code or an Arbitration Agreement to render Orders and Awards, including a sole Arbitrator and all Arbitrators of an arbitration panel.[73]

All administrator rules use the term "persons" or "individuals," when referring to an action taken by an arbitrator. All administrator rules allow parties, upon mutual agreement, to modify any rule. Therefore, it appears parties are free to agree to designate a computer as the arbitrator. To date no administrator has opined on whether or not it would administer a claim where parties have designated a computer to serve as the arbitrator. Modifying administrator rules might best wait until the question of the need for legislation has been resolved.

Conclusion

In the opening paragraphs of this article, the question "Where to begin?" was asked and subsequently answered in the context of feasibility. Here, using the same words, we assume feasibility and now focus on what steps are necessary to start the research ball rolling. The successful creation of a program allowing a computer to act as an arbitrator is closer to becoming a reality than many believe, the concerns raised in this article notwithstanding. The basic technologies needed are rapidly falling into place though they are only part of the equation. The complexities involved go well beyond designing the architecture, curating the data set, and deploying the technology. There are any number of matters that will need to be looked into. Issues such as security and privacy, hacking, the ethics of using a computer to issue a binding award, and legal liabilities involving the provider of such services, are but a few.

The continuing evolution that is producing new and better algorithms and tools for the assembling of data sets, provide hope for capturing and

[73]Forum, Code of Procedures for Resolving Business-to-Business Disputes, Appendix "A", https://www.adrforum.com/assets/resources/Arbitration/Rules/Forum.B2B_Rules.v2.3.pdf.

resolving technical unknowns.[74] Training and testing aside, there are at least three (3) unknowns connected to human behavior:

1. People will need to be persuaded to accept the benefits of a machine-driven system that doesn't apply or even recognize deeply engrained perceptions about what our laws are and how our laws operate.

2. Computer scientists will need to fuse the differences between how humans and computers perceive dispute resolution.

3. And developers will have to grapple with the question: "Will human beings be willing to ever accept the judgment of a machine that knows nothing about gut feelings?"

With all the problems and challenges ahead, there is still no doubt that the effort will yield benefits reaching far beyond the present applications of arbitration as a means to resolving disputes. The well-known maxim "Justice delayed is justice denied," applies to localities where the local judicial system is clogged, corrupt, or otherwise non-responsive to the demands being made. Arbitration is a recognized alternative means for delivering a timely and efficient resolution of disputes no matter the size or complexity. Arbitration by A.I. has the potential to quickly move the benefits of arbitration substantially ahead. Governments at all levels are tasked with providing a judicial system. With few exceptions, there are no limitations on demands for access and yet resources are sparingly provided by legislatures resulting in clogged calendars and over-worked judges and supporting staff. The structures of the systems and protocols require humans to serve other humans thereby adding pressures flowing from the personal needs of all involved and leading to delays brought on by scheduling conflicts and even a failure to have available adequate court room facilities. Frustrated by delay, some stoop to corruption while others take the law into their own hands. No matter the basis for the ill, the respect for law is undermined.

[74]Himabindu Lakkaraju et al., *Identifying Unknown Unknowns in the Open World: Representations and Policies for Guided Exploration* (Oct. 28, 2016), https://arxiv.org/abs/1610.09064.

Without doubt, there is room and an opportunity for any reasonable program with the potential to overcome these problems. Arbitration using A.I. presents such an opportunity. If successfully implemented, the need for additional court rooms, judges, and staff will be significantly reduced, as will associated costs. Scheduling conflicts will be reduced. Delay times will be reduced as computers can handle hundreds, if not thousands of cases every day, 365 days a year. Corruption can be contained because of the difficulties associated with directly corrupting a computer. While not perfect, the issued awards will provide resolution.

The starting point is recognizing the need for research. Developing the data sets, designing the algorithms, and deploying the models will be challenging, requiring an intense effort, and a financial commitment, though the investment required will be less than the ongoing impact of proceeding solely with human arbitrators. The only limitations are a lack of focus on the topic and the willingness of someone to make the needed financial commitment.