Crowdsourcing Arbitration: Using Technology to Leverage Collective Intelligence

Chen CHEN,[*] Colin RULE[**] & Sharon MATHEW[***]

This article examines the use of crowdsourced dispute resolution as a means of resolving disputes in an unbiased, evaluative fashion. The advantages of aggregated collective intelligence over single evaluators have been demonstrated for centuries, but coordination has always been a challenge. Now technology has significantly improved opportunities for coordination, making crowdsourced online dispute resolution (CODR) mechanisms much more feasible. The authors present a framework for crowdsourced arbitration and then apply it to three case studies: the Community Court at eBay.com, the crowdsourced justice system at the Chinese e-commerce site Taobao, and the Kleros crowdsourced jury system. The advantages and challenges of each system are discussed, along with conclusions around the efficacy and utility of crowdsourced arbitration in other applications moving forward.

1 INTRODUCTION

The "wisdom of the crowd" is not a new concept. In the fourth century B.C.E, Aristotle wrote in his book *Politics* that "it is possible that the many, though not individually good men, … when they come together may be better, not individually but collectively."[1] Since Aristotle made that observation more than two thousand years ago, examples of the accuracy of collective intelligence have been well documented. From a crowd estimating the weight of an ox in the 1800s to guessing the location of a sunken submarine in the 20th century, many studies have established that the average of group predictions is often closer to the truth than most individual guesses. Even more interesting is the fact that diverse groups that include both experts and non-experts, are often more accurate than groups composed only of specialists.

In his definitive book on this phenomenon, *The Wisdom of Crowds: Why the Many Are Smarter Than the Few*, James Surowiecki shares data drawn from a wide

---

[*] Adjunct Professor of Law, South China University of Technology and Assistant Director of the Guangzhou Court of International Arbitration.
[**] CEO of ODR.com, Mediate.com, and Arbitrate.com. Former Director of Online Dispute Resolution at eBay and PayPal.
[***] Law Clerk, Beveridge and Diamond.
[1] Aristotle. *Politics*, Book 3, http://www.perseus.tufts.edu/hopper/text?doc=Perseus%3Atext%3A1999. 01.0058%3Abook%3D3, last visited July 16, 2023.

range of cases to demonstrate how collective wisdom can surpass the judgment of any of the individuals who make up the group, even including the predictions of trained evaluators.[2] The ubiquitous television game show *Who Wants to Be a Millionaire* provides a weekly demonstration of this dynamic. Players are offered "lifelines" when they find themselves at a loss in deciding which of the multiple choice answer is correct. One lifeline called "Ask the Expert" enables the participant to discuss the possible answers by phone with an expert nominated by the participant. Another lifeline called "Ask the Audience" enables the participant to poll the audience as to their opinions. The audience members vote to register the answer they believe is correct, and the percentage of votes for each answer is then shown on the screen. Surprisingly, as Surowiecki notes by reference to the American version of this game show, the latter lifeline was more effective in picking the correct answer; the experts answered correctly 65 percent of the time, while the audience was right 91percent of the time.[3]

Surowiecki acknowledges the skepticism surrounding collective judgment, including that of Charles Mackay (who wrote *Extraordinary Popular Delusions and the Madness of Crowds* in 1841),[4] of Bernard Baruch (the early 20th century investor who said "never follow the crowd"), and of philosophers Henry David Thoreau and Friedrich Nietzsche. But the "madness of crowds" (for example, a crowd egging each other on to an act of violence) or a visionary inventor acting on the courage of their convictions (and ignoring conventional wisdom) represent different phenomena from aggregating independent estimations from a large group of unbiased evaluators. Innovation may depend on perceptive individuals who act on dynamics that most others miss, but justice is about fairness, transparency, process, and impartiality, all of which can all be enhanced by collective intelligence; collective judgment leverages fair and transparent processes to aggregate input from many evaluators, so any the influence of any one biased or partial single evaluator is diluted by the other members of the group.

Surowiecki concludes in his book that, given the right incentives and process design, crowds can often make better decisions than individual experts. But he predicates this conclusion on four conditions – diversity in the group, independence of each decision maker, decentralization that fosters local and specialized knowledge, and aggregation of decisions.[5]

---

[2]   James Surowiecki, *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations* (New York, Doubleday, 2004).

[3]   James Surowiecki, excerpt from *the Wisdom of Crowds*, https://www.penguinrandomhouse.ca/books/175380/the-wisdom-of-crowds-by-james-surowiecki/9780385721707/excerpt, last visited July 16, 2023.

[4]   Https://en.wikipedia.org/wiki/Charles_Mackay_(author), last visited July 16, 2023.

[5]   James Surowiecki, *op. cit.*, fn 2.

2 CROWDSOURCING ARBITRATION

While arbitrators are not tasked with making "predictions" and "guesses," the benefits of collective intelligence described above are still quite commonly leveraged in both domestic and international arbitration. Complex cases are often heard by three or five arbitrators on the presumption that aggregating the opinions of multiple evaluators is often fairer and more accurate than relying on a single opinion from a single evaluator. But taking that concept to the next level, what if a single arbitration case could be heard by a dozen arbitrators, or even twenty arbitrators? The dynamics described by Surowiecki, already leveraged for smaller arbitral panels, could potentially benefit more complex cases by convening even larger panels, reinforcing party confidence in the award's independence, impartiality, diversity, and collective expertise.

One of the primary concerns in working with a sole arbitrator is the risk of conflicts of interest. Maybe there were past relationships between the arbitrator and the parties or the counsel representing one of the parties, or maybe the issues in the arbitration cannot be evaluated in an unbiased way due to the arbitrator's past professional experience. A significant amount of time and effort is spent on conflict checks for arbitrators because a fundamental principle of arbitration is that all arbitrators must be independent and impartial. An additional way to combat the risk of bias in a single arbitrator is to select a panel of arbitrators, because then each arbitrator can only influence a portion of the decision, and presumably any biases they harbor would be mitigated by the presence of their fellow arbitrators on the panel.

Crowdsourced online dispute resolution takes this dynamic to the next level. If three arbitrators mitigate the risk of bias from a single arbitrator, then five arbitrators would do an even better job mitigating that risk. But why not expand our arbitration panels to 11 arbitrators, or 25 arbitrators, or even 99 arbitrators? If a case is put in front of 21 arbitrators then any bias harbored by one of the panelists can only influence 1/21 of the overall decision. The larger the jury the smaller the influence of a single decision maker, which reduces the risks from individual conflicts of interest.

The primary barrier to expanding arbitration panels in this way has been cost and efficiency. Convening such a large panel in-person may be cost prohibitive. Qualified arbitrators are often expensive, so empaneling 10 arbitrators instead of 5 may be rejected outright by the parties paying the bills. Also, it would be extremely complicated to coordinate the participation of 21 separate arbitrators in a traditional arbitration process. The case management for such an in-person process would be costly, as finding a time when all ten arbitrators (as well as the

parties and counsel) are available to meet could take weeks of communication and coordination.

But these challenges are becoming more manageable through technology, which has changed the underlying cost benefit calculations behind crowdsourced processes. It is very clear that technology is rapidly transforming the practice of arbitration. Online Arbitration (or, as it was first labeled by Professor Amy Schmitz in 2008, OArb)[6] is quickly becoming mainstream for the arbitral processes. The Internet had been slowly changing arbitration for two decades, but the pace of this change accelerated markedly during the COVID-19 pandemic. Now studies in the United States indicate that the majority of arbitrations are utilizing technology tools like videoconferencing, online filing and case management, and e-discovery, with exclusively in-person hearings becoming the exception rather than the rule.

Software unlocks efficiencies that enable coordination at a much greater scale. The complexity of convening twenty arbitrators for a single case may seem hopelessly impractical face-to-face, but software can make reasonable online what would be impractical or impossible offline. Software platforms can handle the administrative complexity of assembling a large group of decision makers and organizing their input into a coherent decision. Participants can interact asynchronously online, so it isn't necessary to coordinate 20 separate schedules. OArb has the potential to merge the neutrality and efficiency of a well-designed arbitration process with the efficiency, fairness, and transparency of crowdsourcing.

3   A FRAMEWORK FOR CROWDSOURCED DISPUTE RESOLUTION

The concept of crowdsourced arbitration is not new. The crowdsourced platform iCourthouse.com[7] was nominated for a Webby Award in 2001,[8] and a US patent was filed describing the concept in 2009.[9] An Israeli startup, allrise.com, provided crowdsourced resolutions to online users in 2010,[10] and even the American daytime television program The People's Court[11] put together a

---

6   Amy J. Schmitz, "Arbitration in the Age of COVID: Examining Arbitration's Move Online," Cardozo Journal of Conflict Resolution, vol. 22, no. 2 (2021) p. 245; University of Missouri School of Law Legal Studies Research Paper No. 2020-27. Available at SSRN: https://ssrn.com/abstract=3699778, last visited July 16, 2023.

7   Https://web.archive.org/web/20011011232729/http://www.i-courthouse.com/main.taf?&redir=0, last visited August 21, 2023.

8   Https://winners.webbyawards.com/2001/websites/general-websites/government-law/157582/icourthouse, last visited July 16, 2023.

9   "System and method for community-based dispute resolution," https://patents.google.com/patent/US20110047007A1/en?oq=20110047007, last visited July 16, 2023.

10   Https://web.archive.org/web/20101114042921/http://www.allrise.com/, last visited July 16, 2023.

11   Https://en.wikipedia.org/wiki/The_People%27s_Court, last visited July 16, 2023.

website, peoplesourtraw.com,[12] that leveraged similar approaches to resolve a wide variety of disputes using video testimonials. Gao Wei classified these current crowdsourced dispute resolution mechanisms into four categories: (1) online opinion polls; (2) online mock trials; (3) crowdsourced procedures rendering decisions that are enforceable by private authorities; and (4) decentralized crowdsourced mechanisms using blockchain technology.[13] New experiments fitting into these categories crop up on a regular basis.

In 2017 Daniel Dimov wrote his doctoral thesis on crowdsourced online dispute resolution (which he called CODR).[14] His framework for considering the design of crowdsourced dispute resolution systems crystalized the key considerations that need to be factored into the design of an ODR process that leverages collective intelligence.

Dimov lays out four main building blocks: 1) the crowd, 2) the incentives, 3) the types of disputes addressed, and 4) the process. He then lays out several criteria that must be considered in designing a CODR system: the mechanism used for resolving disputes, the eligibility to participate in CODR as a member of the crowd, the number of members of the crowd, composition of the third neutral party, and the use of deliberations between the members of the crowd.

We will now share three case studies of crowdsourced online dispute resolution projects (eBay, Taobao, and Kleros), examine their building blocks (e.g. the crowd, the incentives, and the types of disputes), and apply Dimov's criteria to distill lessons learned.

3.1   CASE STUDY #1: EBAY'S COMMUNITY COURT

In 2008 the online dispute resolution team at eBay created the *Community Court*, a crowdsourced platform for resolving e-commerce disputes.[15] The design was simple: if a seller on eBay felt that they had received a bad review that they did not deserve, they could log into the Community Court platform and explain why the feedback should be removed. To support their case the seller could upload images, text, or other digital files relevant to the issue. Once the case was filed, the

---

[12]   Https://web.archive.org/web/20080905143045/http://www.peoplescourtraw.com/, last visited July 16, 2023.

[13]   Wei Gao, "Let the collective intelligence shine through," Peking University Law Journal, vol. 6, no. 2 (2018) pp. 283-304, https://www.tandfonline.com/doi/abs/10.1080/20517483.2018.1603645, last visited July 16, 2023.

[14]   Daniel Dimov, *Crowdsourced Online Dispute Resolution*, Leiden University Center for Law and Digital Technologies, SIKS Dissertation Series No. 2017-17, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3003815, last visited July 16, 2023.

[15]   Colin Rule and Chittu Nagarajan, "Leveraging the Wisdom of Crowds: The eBay Community Court and the Future of Online Dispute Resolution," ACResolution Magazine, Winter (2010) pp. 4-7, https://colinrule.com/writing/acr2010.pdf, last visited July 16, 2023.

Community Court platform automatically contacted the buyer to offer them an opportunity to respond. The buyer had the benefit of seeing the seller's submissions, and they could then make their case and provide relevant additional information as well. Once the buyer's submission was complete, the seller was provided one final opportunity to rebut the buyer's assertions in text only.

After the case information was submitted by both parties the Community Court presented the information to a randomly-selected panel of jurors. Jurors in the Community Court were eBay members who had previously applied to serve and who met the stringent eligibility criteria (e.g. a significant period of time on the eBay site, a positive feedback rating, and no transaction history with either the buyer or seller involved in the case). Each juror reviewed the information submitted by the seller and buyer in its entirety before entering their decision. The outcome options for jurors were: agree with the buyer, agree with the seller, or unable to make a decision.

Each case in the eBay Community Court was heard by 21 jurors. If more than half of the jurors agreed with the seller then the case was decided in the seller's favor and the feedback was removed from eBay's reputation system. If more than half of the jurors disagreed with the seller then the feedback stood as left by the buyer. The Community Court leveraged hundreds of jurors to resolve thousands of cases over several years of availability.

The team administering the Community Court quickly learned the strengths and weaknesses of the CODR model and moved rapidly to refine the design in response. The team also partnered with several academic institutions to analyze the performance of the platform, including Pepperdine University, Creighton University, and Harvard Law School. Researchers from those institutions conducted surveys with the buyers, sellers, and jurors who used the platform, as well as gathering feedback from eBay users who were familiar with the concept but had never used the platform themselves.

One element that proved to be crucial was case assignment. When a new filing was received in the Community Court the system did not immediately assign the case to 21 jurors on the approved panel. Instead, the Community Court assigned cases to jurors on a first-come first-serve basis (meaning the next 21 qualified jurors that logged in would be notified of the case available for them to review). Once they accepted the assignment, the jurors only had access to the case for a limited period of time, so they could not log out of the platform and come back to that particular case. This constraint ensured that jurors could not contact the involved buyers or sellers or gather inappropriate information.

The initial design hypotheses held that each Community Court jury should be balanced between buyers and sellers for fairness because there was a concern that buyers would be more likely to decide in favor of buyers and vice versa. What the

team found, however, is that seller-jurors were often harder on their fellow sellers than buyer-jurors. This is likely because sellers understood their obligations better than buyers, so sellers could more accurately assess the performance of other sellers against those obligations. In addition, seller-jurors understood that bad behavior from any individual seller reflected negatively on the overall community of sellers, so there was an incentive to maintain high standards.

The original concern was that there would not be many applicants to be jurors, but in fact there were more than enough applications to support the case volume. Jurors were not paid for their services, but the Community Court team budgeted for small incentives to be provided to jurors to thank them for their work. However, no incentive payouts were necessary because the jurors were more than willing to participate out of their sense of service to the community (this may be a particular feature of the eBay user environment, which was very engaged, so other online sites will need to determine whether their crowdsourced implementations would demonstrate similar commitment from community members.)

The Community Court team also developed several models to monitor jury verdicts and identify troublesome patterns. These models monitored how many times jurors were in the minority on a decision, how long they took to review the information submitted by the buyer and seller, and the rationale they provided to back up their decisions. If a juror displayed any concerning patterns, the team could refer to them cases that had already reached an outcome as a test, or the platform could stop referring cases to them altogether.

The Community Court won several innovation awards at eBay and PayPal, and eBay filed for a patent on collective dispute resolution ("System and Method for Community Based Dispute Resolution") which was granted in 2011. However the patent is limited to eCommerce, and as such does not cover the application of crowdsourcing to arbitration.[16]

Applying Dimov's framework to the eBay Community Court, we can identify the *crowd* (the pool of eBay users who volunteered to serve as jurors) the *incentives* (a drive to contribute to trust and safety within the eBay community), the *types of disputes addressed* (feedback/reputation disputes between sellers and buyers) and the *process* (each party submitting their perspective, the jurors evaluating the information and rendering a decision, and eBay enforcing the outcome). The application of Dimov's framework also illustrates the high bar for panelist eligibility, the relatively small panel size, and the lack of deliberation between the panelists in the Community Court design.

---

[16] "System and Method for Community Based Dispute Resolution," US Patent #20110047007-A1, http://ombuds.org/files/crowdpatent.pdf, last visited July 16, 2023.

3.2   CASE STUDY #2: TAOBAO

Founded in 2003, Taobao.com is one of the biggest online retail platforms in China, with approximately 70% market share. It has nearly 500 million registered users and over 60 million daily visits. More than 48,000 items are sold every minute on the platform. For the fiscal year of 2021, the gross merchandise volume of Taobao's ecosystem exceeded $1.17 trillion, with the sales of Taobao's *Double 11 Shopping Festival*, the counterpart of the Black Friday sales in the U.S., reaching $78 billion alone.

Ever since its launch, the volume of transactions on the Taobao platform have given rise to a large number of disputes that require quick, efficient, low cost resolution. In 2012, in an effort to better sustain and profit from the business activities on the platform, Taobao established a crowdsourced online dispute resolution (CODR) system, providing a blueprint for other e-commerce platforms in China facing similar issues. The effectiveness of Taobao's CODR system is vitally important considering the volume of sales generated on each online retail platforms and the profound impact of e-commerce on China's economy as a whole.

In view of the sheer volume of disputes arising from its online retail business, and inspired by the Resolution Center mechanism created by eBay several years earlier, Taobao set up an evaluation center in 2011 to facilitate the timely resolution of disputes. In that year Taobao received approximately 12.6 million complaints by phone call, and it processed more than 2.16 million complaints related to consumer rights. Most of the large number of complaints could be categorized into groups, such as delayed package delivery or inferior quality of goods. The amounts in dispute involved were often small. It was a matter of critical importance for Taobao to address these issues effectively considering the great cost of processing these large complaint volumes, not to mention the consequential impact if even a portion of those complainants decided to take their cases to the courts. As a result, building an efficient dispute resolution system was on top of Taobao's priority list.

Taobao's first attempt was to create the role of "salesclerk." A salesclerk is essentially a designated customer service representative assigned to a case whenever a complaint arises, acting as a facilitator for the resolution of the dispute. Under Taobao's dispute resolution mechanism, only the buyer can initiate the process. Once the case is filed and evidence is exchanged, the salesclerk then renders a decision (though one or both parties must have requested salesclerk involvement). This mechanism put substantial effort into facilitating settlement by mutual agreement prior to an imposed decision. If the complainant does not take action for a specific time after the submission of a dispute resolution application, it is

deemed that a settlement has been reached and the case is automatically closed. If a decision is to be rendered, before doing so, the salesclerk gives the parties 24 hours as the last opportunity to settle. The salesclerk mechanism served its intended purpose, but it also had serious drawbacks, such as lack of supervision and unfair decisions. Accordingly, salesclerks gradually lost some of their credibility with users.[17]

In 2012, under the slogan "let collective intelligence shine through," Taobao established its CODR mechanism as an alternative to the salesclerk mechanism. The CODR mechanism, first of its kind in China, was designed to resolve the following types of disputes: (1) seller violations, including but not limited to violations discovered by Taobao's active investigation or complaints/reports by buyers, as well as seller appeals after the violations have been processed; (2) transaction disputes, such as the right to reimbursement, payment, or compensation; (3) other kinds of disputes or services deemed necessary or appropriate for CODR jurors, such as consumer rights protection, community policy decision-making, or identification of a malicious product review.

Jurors became the critical component of Taobao's CODR mechanism. A Taobao member must meet certain criteria before becoming a CODR juror: (1) be a Taobao member for at least a year and hold a membership level over VIP2;[18] (2) have a Sesame Credit score[19] of at least 600; (3) have at least a Taobao credit score shown as three hearts; and (4) be involved in less than three transaction disputes that require the intervention of Taobao.[20] For sellers the criteria are even more demanding, which additionally require a higher Taobao credit score, a lower refund rate than average, and no violation that is considered as serious and intentional by the standards of Taobao's regulations within 6 months after becoming a CODR juror. There are also incentives for the juror to improve their skill,[21] such as improving the ranking of the juror's account, or accumulating scores related to charitable activities. According to published statistics, Taobao's CODR jurors' ages range from 16 to 74, the ratio of male jurors to female jurors is

---

[17]   Xiang Zhou, "Description and Explanation: Dispute Resolution Mechanism of Taobao – China's Experience and Observation of ODR," Journal of Shanghai Jiao Tong University (Philosophy and Social Sciences), vol. 4 (2021).

[18]   Taobao has a credit system for its VIP users, ranking from VIP0 to VIP6, with higher number indicating more senior status.

[19]   The Sesame Credit score, designed by an independent third-party credit agency under Ant Group, objectively presents an individual's credit status through cloud computing, machine learning and other technologies. It has been used in credit cards, consumer finance, financial leasing, hotels, rentals, and other scenarios in China.

[20]   Haitao Shang, "Analysis of folk norms and operation mechanism in crowdsourcing dispute resolution," Folk Law, vol. 2 (2020).

[21]   Jurors can improve their skills by completing online simulation tasks designed by Taobao.

1.5:1, and Zhejiang, Guangdong, and Jiangsu are the three provinces with the most registered jurors.

Taobao assigns cases randomly to its jurors to render a verdict. Under its 2013 CODR regulations, within the given time limit, if there are less than 31 jurors who vote for a particular decision, then the verdict is invalid. A party who receives at least 16 votes will have won the case. In 2014, the rules were changed to provide that, within the given time limit, if either side fails to receive 16 votes, the final verdict is invalid. This new rule removed the requirement for the number of jurors needed for a valid verdict. Additionally, the time limit for the final verdict was extended from 24 hours to 48 hours, then to 168 hours. All these changes indicate Taobao is tuning its rules towards more valid verdicts, and actively promoting and refining its CODR system.

By 2018, Taobao had 4.31 million jurors, of whom 1.72 million participated in the resolution of disputes that had reached a valid verdict. The total number of valid verdicts exceeded 100 million. The consensus of the user community was that it successfully provided an efficient, fast, low-cost means of resolving certain types of e-commerce disputes. However, for a platform like Taobao, with such extensive influence, Chinese internet users and resources, continuous improvement is essential not only for the development of Taobao but also for the benefit of the public as a whole.

Applying Dimov's framework to Taobao's CODR system, we can identify the *crowd* (the pool of Taobao users who volunteer to serve as jurors) the *incentives* (a drive to contribute to trust and safety within the Taobao community), the *types of disputes* addressed (seller violations and appeals, transaction disputes, and community policy decision-making) and the *process* (each party submitting their perspective, the jurors evaluating the information and rendering a decision, and Taobao enforcing the outcome). The application of Dimov's framework also highlights the high bar for panelist eligibility, the larger panel size, and the participation threshold for decision enforceability.

3.3   CASE STUDY #3: KLEROS.IO

The CODR systems developed by Taobao and eBay pioneered the application of collective intelligence to e-commerce disputes, acknowledging the high transaction cost of conventional litigation and the appeal of outsourcing decision making to a community of jurors as opposed to a single arbitrator. Even so, these systems were limited to commercial transactions that took place on these respective websites. In this case study, examining Kleros (a blockchain-based dispute resolution system) we consider a CODR system not limited to disputes on a

specific platform but applicable to agreements brought into effect through smart contracts.[22]

The first stage of the Kleros process involves drafting a smart contract between the two parties. The contract must have a specific clause that mandates arbitration by Kleros along with specific details on the Kleros court to which the parties will be subject in the event of a dispute. Kleros has a court hierarchical system with eight sub-courts[23] and a general court at the top where all appeals are heard. Once the clause has been triggered, the disputed amount will be held in escrow until it is resolved. The parties will also each deposit an arbitration fee, an amount that varies based on the court and on the complexity of the issue.

The second stage of the Kleros process involves securing the evidence. Blockchains (online ledgers that are almost impossible to tamper with) are used to handle and secure the evidence, providing reassurance to users that the evidence will be incorruptible. Once the evidence has been secured, the jury is selected to evaluate the dispute. Each court specifies the set of skills required from a juror along with the minimum "stake" they must bid (using a special online token issued by Kleros called a Pinakion) to be selected as a juror. The higher the juror's bid, the greater the likelihood that they will be selected as a juror. The juror's identity is not revealed to protect them from "intimidation or retaliation."[24] These two features of incorruptibility of evidence and anonymity of jurors ensures that the dispute resolution will be fair.

Once the jury has been selected, they are given access to the evidence and are required to vote. Kleros uses an incentive mechanism based on the Schellingcoin,[25] where jurors who vote coherently with the majority will be rewarded, and those who vote in the minority will lose a portion of what they had staked.[26] Jurors have

---

[22] Smart contracts are a blockchain based agreement where all or part of an agreement is automatically executed. This form of contracting is preferred over traditional contracts because of the "security, permanence and immutability" that they offer. *See* Alex Lipton and Stuart Levi, "An Introduction to Smart Contracts and Their Potential and Inherent Limitations," The Harvard Law School Forum on Corporate Governance (May 26, 2018), https://corpgov.law.harvard.edu/2018/05/26/an-introduction-to-smart-contracts-and-their-potential-and-inherent-limitations/, last visited September 17, 2023.

[23] This includes the Blockchain (Technical and Non-Technical Subcourt), the Marketing Services Subcourt, the English Language Subcourt (disputes involving the quality of written content), the Video Production Subcourt, the Curation Subcourt, the Data Analysis Subcourt, the *Corte General en Espanol* (disputes that require a high level of Spanish proficiency to be resolved) and the Humanity Court (proof of human disputes), *see* https://court.kleros.io/courts, last visited July 16, 2023.

[24] Kleros, "Kleros FAQ," https://docs.kleros.io/kleros-faq, last visited September 17, 2023.

[25] A theory developed by game theorist, Thomas Schelling, where he provides a specific set of facts to random people and states the people arrive at a focal point (Schelling point) in the absence of communication and in the presence of incentives for coherent decisions. *See* Thomas C. Schelling, *The Strategy of Conflict* (Harvard University Press, 1981), https://www.hup.harvard.edu/catalog.php?isbn=9780674840317, last visited July 16, 2023.

[26] Clément Lesaege, Federico Ast, and William George, "Kleros Short Paper v1.0.7," https://kleros.io/whitepaper.pdf (last visited September 17, 2023).

no way of interacting with each other or colluding to game the outcome. In the absence of this, they are incentivized to vote for the decision that they assume the majority of other jurors will reach when given the same evidence. Kleros, therefore, incentivizes truth seeking and eliminates negative occurrences such as bribery or collusion that are prevalent in traditional courts or conventional commercial arbitration.

Parties may either accept the majority vote (decision) of the jury or appeal the decision to a larger jury. However, it is relevant to note that with each appeal, the arbitration fee increases steeply, since the arbitration fee is based on the number of jurors. Once both parties have accepted the final decision, the disputed amount placed in escrow at the beginning of the process is transferred to the winning party, and the arbitration fees are distributed amongst the jurors based on how they voted.

All transactions on Kleros take place using a special cryptographic token called Pinakion (PNK). Kleros' management of this token helps to protect the system against attacks and to prevent parties from gaming the dispute resolution process. As an example, if a potential juror acquires a majority (51%+) of the staked amount in a given case, and the minority-holding community that believes that an attack is taking place, the community may collectively remove the PNK staked by the attacker from the total pool of PNK and create a new system absent the attacker's PNK holding.[27]

Kleros offers numerous benefits to dispute resolution, the first being its efficiency. The Kleros design recognizes the globalized nature of transactions and is able to provide an arbitration forum where small claims between clients in one country and service providers in another country can be resolved without needing to consider complex questions of jurisdiction. At present, there is no traditional court system that is designed to resolve such disputes. Further, each Kleros court has a stipulated session period or time-period within which the jurors must either submit their vote or lose a portion of the PNK they had staked. This ensures that disputes are resolved in a timely manner.

Second, Kleros seeks to balance any economic inequalities between parties through its appeal process. In a traditional court system, parties that have greater resources can dictate the course of the dispute by using the appeal process inherent in the system. While Kleros also has an appeal process, it safeguards against manipulation by making the process exponentially more expensive with each appeal and by allowing third parties to "crowdfund" the appeal fees for parties who cannot afford it. This means that anyone can contribute to a portion of either

---

27   Kleros, "PNK Token," https://docs.kleros.io/pnk-token (last visited September 17, 2023).

parties' arbitration fees and these crowd-funders will ultimately be financially rewarded if their party wins.[28]

Third, Kleros offers transparency to users involved in the dispute resolution process. Some of the challenges in traditional court systems are that parties may tamper with the evidence, submit false evidence, or even manipulate decision makers such as jurors or judges in order to tip the scales in their favor. As mentioned above, Kleros seeks to address these challenges through blockchain technology which records transactions in a verifiable, secure, and permanent way. This technology also maintains the anonymity of jurors in the dispute. As a consequence, neither the parties nor individual jurors have any way of contacting another juror, making it nearly impossible for any single actor to influence the decision-making process.

Fourth, Kleros ensures that jurors do not communicate with each other in the decision-making process and ensures the independence required to generate the best decisions. It has no barriers to entry in terms of the kinds of people who can participate as jurors and therefore, enables localized and specialized knowledge from jurors across the world. And lastly, it provides incentives in the form of PNK to jurors who vote coherently, thus incentivizing independent jurors to vote truthfully.[29] Applying Dimov's framework to Kleros' CODR system, we can identify the *crowd* (the individuals who bought PNK and staked their participation) the *incentives* (an interest in contributing to online fairness and participating in crowdsourced justice), the *types of disputes* addressed (smart contract and virtual currency disputes) and the *process* (evaluating information, rendering decisions, and financial incentives for agreeing with the majority of other jurors). The application of Dimov's framework also illustrates the sophisticated tools to prevent participant manipulation, the appeal layers, and the transparency.

4 COMBATTING ABUSE

Is it possible to manipulate a crowdsourced arbitration process? Certainly, but with careful systems design, mechanisms can be put in place to make such manipulation far more difficult. If a case is sent to 200 jurors and those jurors have multiple weeks during which they can evaluate the facts submitted in the case, that creates a window where an agent looking to influence the process might try to identify and

---

[28]   Kleros, "Introduction to Kleros," https://kleros.gitbook.io/docs/kleros-faq, last visited July 16, 2023.
[29]   This is based on the idea that jurors who don't know each other will vote for the true answer because they expect others to vote for the true answer if they want to be part of the majority. *See* Clément Lesaege, Federico Ast, and William George, *op. cit.*, fn 26.

contact a large number of the jurors so as to maximize the chance that the decision would go in their favor. However, if the crowdsourced arbitration process is designed so that the jurors are kept anonymous, they are tested to ensure they have never had any interaction with the parties in the case, and they are given a limited amount of time to evaluate the submitted materials and render their decision, the window for influence is much smaller.

With appropriate technology it is very easy to aggregate large amounts of data and then utilize that data to identify troublesome patterns. Jurors who exhibit behavior that may call into question the fairness and thoroughness of their decision-making can be demoted in the juror selection process for future cases. Perhaps the time each juror spends looking at the information submitted in the case, and the percentage of the time that that juror agrees with the majority of their fellow jurors, can be used to generate a juror trust score. That score can then be used by a case manager to identify the best performing jurors and to steer cases in the direction of those jurors over time.

One challenge is adequately ensuring the diversity of jurors. Maximizing confidentiality may be a design priority (such as Kleros' use of anonymity) but it can preclude administrators from knowing the identity of jurors, which makes ensuring demographic diversity very difficult. Federico Ast, Kleros' CEO, has reported that a majority of the jurors on Kleros are young adults or students who possess NFTs, who were comfortable with Web3,[30] or who have a reasonable understanding of blockchain technology.[31] In a self-selected pool of jurors such as that it may be difficult to eliminate inherent biases that exist in groups of like-minded people, therefore, resulting in bad or biased judgments over time.

Designers of CODR systems need to make tradeoffs between competing values. CODR systems can be designed to keep case data confidential (like Taobao and eBay) or release it publicly to underscore transparency (like Kleros). But sharing data publicly may create opportunities for juror bias. An essay by William George, Director of Research at Kleros, revealed how such biases might arise from case transparency.[32] He explains that in a dispute between a small business owner and a freelancer, if jurors are given access to the website of the small business owner, they might access additional information about the business not provided as evidence (such as Yelp reviews). This information might sway

---

[30]   An internet platform built using decentralized blockchains. *See* Kevin Roose, "What Is Web3?," The New York Times (March 18, 2022), https://www.nytimes.com/interactive/2022/03/18/technology/web3-definition-internet.html, last visited September 17, 2023.

[31]   Interview with Federico Ast, CEO of Kleros, on December 15, 2022.

[32]   William George, "Kleros and Mob Justice: Can the Wisdom of the Crowd Go Wrong?" in *Dispute Resolution: The Handbook of Decentralized Justice* (2020) p. 134, https://kleros.io/book.pdf, last visited July 16, 2023.

jurors either for or against the small business owner. Another tradeoff may be between financial incentives for jurors. Kleros rewards jurors who stake larger amounts of PNK tokens with more cases, which incentivizes quality juror service, but it may also lead to richer jurors (those with more money to risk) getting selected more often than poorer jurors.

5   CODR AND ARTIFICIAL INTELLIGENCE

Some might assume that the rise of machine learning and artificial intelligence will obviate the need for crowdsourced dispute resolution mechanisms because we will be able to create algorithmic judges that will have their own mechanisms to ensure their decisions are fair and unbiased.[33] However, the benefits of CODR are not limited only to human powered processes. Instead of one all-seeing algorithm deciding all types of arbitration cases, it is more likely that we will see the emergence of many different machine learning powered evaluative algorithms, and we may decide to crowdsource algorithmic neutrals in much the same way we crowdsource human jurors.

Given that machine learning-powered virtual arbitrators will never tire, and will work 24/7 for a fraction of a penny's worth of electricity, we may see the emergence of panels of robot judges behaving much the same way that human jurors do in CODR processes. Even for a low dollar value case it might be possible to receive an evaluation from 15 or 25 different algorithmic evaluators, and whichever party gets a majority of the algorithmic evaluators to agree with them will win the case. These algorithmic evaluations must always be subject, of course, to human review upon the request of either party, but as time goes on (and the accuracy and quality of AI evaluators improves) the requests for human review will likely become less common.

This dynamic echoes the early days of optical character recognition (OCR) technology. There were several different OCR algorithms created around the same time in the 1990s but none of the algorithms on their own could get above a 97-98% accuracy rate, which is too low for effective optical character recognition. However, one technologist had the idea of using all three of the major OCR Technologies on the same document and then aggregating their results. The combination of the three separate algorithms was able to achieve accuracy above 99%, which made OCR useful. As John Handey and Thomas Hickey explained at the time, "outputs from several OCR systems can be combined to produce a more

---

[33]   Gizem Halis Kasap, "Can Artificial Intelligence ('AI') Replace Human Arbitrators? Technological Concerns and Legal Implications," Journal of Dispute Resolution, vol. 2021, no. 2 (2021), https://ssrn.com/abstract=3863136, last visited September 18, 2023.

accurate result."[34] The same may be true for algorithmic evaluators leveraging AI and machine learning.

6   CONCLUSION

The lines between traditional arbitration and CODR mechanisms are beginning to blur. In May 2021, a Mexican court gave validity to an agreement where the arbitration clause specified that the arbitrator must use CODR (specifically Kleros) to govern the arbitral decision.[35] The court did not directly enforce the Kleros award but gave validity to the clause that required the arbitrator to incorporate the Kleros award into the traditional arbitral award.[36] Even so, this is a big win for CODR services such as Kleros and could lead to an expansion of Kleros' caseload into other areas. It is likely that this case will be followed with similar decisions in other jurisdictions reaffirming outcomes from CODR processes.

One of the most contentious areas in social media is content regulation. Tech giants like Facebook and Twitter have faced growing criticism for taking down hundreds of accounts, such as in India where posters are criticizing the government for their management of farmer protests.[37] There is an increasing distrust of internal mechanisms and discretion being left in the hands of big tech in regulating content. CODR systems could be integrated into these platforms, not only to serve as a transparent external party in regulating content, but also to facilitate power being put back into the hands of the users in regulating their own communities. This may be the next area of expansion for CODR approaches.

Arbitration has developed its own robust sets of rules and standard procedures over the last hundred years, and those designs have served as the foundation for

---

34   John C. Handey and Thomas B. Hickey, "Merging optical character recognition outputs for improved accuracy," Computer-Assisted Information Retrieval (*Recherche d'Information et ses Applications*)" – RIAO 1991, 3rd International Conference, *Universitad Autonoma de Barcelona*, Spain, April 2 – May 5, 1991, https://www.researchgate.net/publication/221509893_Merging_optical_character_recognition_outputs_for_improved_accuracy, last visited July 16, 2023.

35   Mauricio Virues Carrera, "Accommodating Kleros as a Decentralised Dispute Resolution Tool for Civil Justice Systems: Theoretical Model and Case of Application," Annex VIII, https://ipfs.kleros.io/ipfs/QmfNrgSVE9bb17KzEVFoGf4KKA1Ekaht7ioLjYzheZ6prE/Accommodating%20Kleros%20as%20a%20Decentralized%20Dispute%20Resolution%20Tool%20for%20Civil%20Justice%20Systems%20-%20Theoretical%20Model%20and%20Case%20of%20Application%20-%20Mauricio%20Virues%20-%20Kleros%20Fellowship%20of%20Justice.pdf, last visited July 16, 2023.

36   Maxime Chevalier, "Arbitration Tech Toolbox: Is a Mexican Court Decision the First Stone to Bridging the Blockchain Arbitral Order with National Legal Orders?," Kluwer Arbitration Blog, March 4, 2022, https://arbitrationblog.kluwerarbitration.com/2022/03/04/arbitration-tech-toolbox-is-a-mexican-court-decision-the-first-stone-to-bridging-the-blockchain-arbitral-order-with-national-legal-orders/, last visited July 16, 2023.

37   Karan Deep Singh, "Twitter Blocks Accounts in India as Modi Pressures Social Media," New York Times, February 10, 2021, https://www.nytimes.com/2021/02/10/technology/india-twitter.html, last visited July 16, 2023.

arbitration's growth around the world, both domestically and internationally. But the emergence of technology and the digitation of society is disrupting many previously entrenched institutions, and arbitration is no different. Geographic assumptions around the importance of jurisdiction, site, and venue that made sense fifty years ago may no longer be true in a world where parties transact over the internet. Eventually predicating procedural rules based on location may begin to feel like an anachronism.

We should always be open to questioning the core tenets of arbitration process design, especially in light of the new capabilities that technology is bringing to the table. For example, preventing arbitrators from contacting each other during the evaluation process (as some crowdsourced processes dictate) is a radical departure from current arbitration practice. But perhaps questioning some of our long-held practices is a healthy exercise. If we merely use technological advances to administer the same arbitration processes that were being conducted 50 years ago, just faster and more efficiently, we may be missing out on opportunities to expand the reach of arbitration. Questioning our assumptions and rethinking our methodologies can help us avoid merely "paving the cow paths"[38] of our previous practice.

An openness to the benefits of crowdsourced arbitration is a perfect case study for how we can embrace the opportunities that may lie ahead. If our CODR systems, following Surowiecki's advice, exhibit diversity in the arbitrator pool, independence of each arbitrator, decentralization to foster specialized knowledge, and effective aggregation of decisions, crowdsourced processes offer great potential for expanding the reach and effectiveness of arbitration over the coming years.

---

[38]   Jim Highsmith, "Paving Cow Paths," Agile Connection, June 21, 2005, https://www.agileconnection.com/article/paving-cow-paths, last visited September 8, 2023.