Should legal disputes be determined by artificial, rather than human, means?

The modern legal environment is convulsing with uncertainty as a potential means of dispute resolution is gaining support: AI. Due to an enormous backlog of cases, multiple countries are aiming to use artificial intelligence to replace or supplement human decision making. For instance, in 2017, an “cyber court”¹ was set up in Hangzhou, China. Using an online platform, litigants appear by video as a black-robed virtual judge hears their case. This court handles online trade disputes, copyright cases, and e-commerce product liability claims. However, despite the appeal of artificial dispute resolution, I firmly believe that it has a destructive potential and should not be employed at all. In the first part of my essay I will argue that if artificial intelligence becomes an adjudicator, it will shift the values surrounding legal systems, pushing towards an “algorithmic” framework. Then I will explore the impact this shift may have on society, namely its potential to alienate the public from the law. In the final part I will discuss whether it is desirable to combine human and artificial adjudication.

Technology is a formidable force. It has the potential to alter not only the way in which the law applies its rules, but also its core values. Most legal systems are characterized by a mixture of equitable and codified decision-making, with an emphasis on the former, especially in more complex fields (for instance criminal or contract law). I shall explain equitable justice first, and then move on to codified justice.

As Aristotle put it, equity² is a form of justice that prefers to adapt general rules to particular cases. Therefore, the decision-making which is based on it involves case-specific judgement and moral reasoning, preferring individual discretion to standardization. As an example, criminal law judges take into account specific mitigating factors (such as age, substance abuse) when they decide the defendant’s fate, instead of relying on a black-and-white rule.

In contrast, codified decision-making prefers to apply a standardized procedure to a set of facts. It aspires to frame general rules in advance to provide for any case the future may bring. Therefore, it pays little heed to the unique facts and circumstances discovered in individual proceedings. An example of codified decision-making is the federal sentencing guideline, which makes use of a rigid system of points to determine the sentence that the perpetrator must serve.

Having explained these concepts, it becomes clear that, if artificial intelligence were to resolve disputes, it would employ the “codified”, algorithmic model of decision-making. This would present several advantages, namely heightened efficiency, transparency and, presumably, minimized arbitrariness.

With regards to efficiency, an AI mediation system has a unique capacity for mass deployment, being able to resolve an indefinite number of cases in a short span of time. This would ease the

backlog of cases, an aspect which is appealing to overburdened legal systems. In fact, the Estonian Minister of Justice\(^3\) announced that a robot judge was being designed to adjudicate a number of small claims disputes, so as to clear their backlog of cases. Furthermore, the capacity for mass deployment of AI judges can render litigation more accessible and cheaper to the public, resulting in a “democratization of justice”.

Moreover, resolving a dispute by artificial means can ostensibly lead to a more transparent verdict. Machine learning systems do not take into account the variables which fail to accurately predict outcomes\(^4\). This feature is in stark contrast to humans, who may lie or not even realize the factors that lead to their decision to, for instance, discredit the reliability of a witness. It is also easier to probe algorithms for bias. But this can only be done by experts, meaning that, despite the previous claims, the potential bias in an AI-given verdict remains just as veiled to the public as it can be in a verdict given by a human judge.

As for arbitrariness, it is claimed that AI is devoid of the discrimination and cognitive bias displayed by humans. Nonetheless, this point is highly contentious, because human bias plays a large part in creating adjudicatory algorithms. This is because any algorithm needs to be coded by a human, and thus stands a high chance of “absorbing” bias\(^5\). For instance, imagine that a programmer is trying to create an AI program that predicts the probability that the individual will reoffend. If the programmer harbors any prejudice against people of color, this will carry over into the program itself, resulting in the prediction of a higher rate of reoffending among the aforementioned group. This was the case in a criminal justice algorithm employed in Florida, which mislabelled African-American defendants as high-risk at almost twice the rate it mislabelled white defendants\(^6\).

Despite a number of advantages, the justice that AI would bring about raises formidable issues. Firstly, a codified, artificial intelligence-based system is inherently rigid. Artificial intelligence makes use of programmed algorithms to process information and reach a predetermined outcome\(^7\). Therefore, it would have difficulty in dealing with those aspects of the law which require discretion, such as questions of interpretation where legislation is unclear or where matters of morality are involved. An exemplification of the latter would be the definition that murder occurs only when someone kills somebody with a “mens rea”. How could an AI judge, devoid of notions of morality or equity, establish whether murder had been committed?

The inability of a robot judge to use discretion has a significant corollary. Namely, as noted by Justice Melissa Perry\(^8\), the resolution of disputes by artificial means will call for the replacement of discretionary laws with more objective, black-and-white ones. For instance, there could be a phasing out of “mens rea” as the basis of guilt in criminal law, with it being replaced by objective criteria to determine whether conduct is criminal. But black-and-white laws are not

\(^3\)https://futurism.com/the-byte/estonia-robot-judge
\(^5\) idem 4)
\(^6\)https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing
\(^7\)https://www.sas.com/en_us/insights/analytics/what-is-artificial-intelligence.html
desirable, for they can result in decisions which lack the necessary nuance and individualization to mete out justice.

Secondly, human adjudicators play a significant role in the community, which extends beyond their act of resolving disputes. It cannot be denied that their commentary serves both a professional and a social purpose. In the professional sense, it informs lawyers about a way of approaching legal issues, which can improve their skills in the future. In the social sense, it yields valuable civic knowledge. Judicial commentary also heightens the transparency of the legal system, which leads to an increased acceptance of the law.

But it stands to reason that AI dispute resolution could render judicial commentary obsolete. Codified decision-making, by nature, tends to eliminate the need for any explanation or legitimation apart from the compliance with a set of standardized procedures. Moreover, robots do not “think” in the same way as humans (for instance, they do not make the same causal inferences). Therefore, it would be impossible for them to hand out a verdict which includes the genuine commentary and reasoning used by human judges. This change would deprive the community of the aforementioned benefits that judicial commentary yields.

The second aspect that erodes the appeal of AI adjudication is that it will insulate the legal system from the public. As noted by Justice Benjamin Cardozo, “the final cause of the law is the welfare of society”. Yet AI adjudication goes against this principle, opening up a chasm between society and the legal system.

The first aspect that weakens the appeal of artificial decision-making to the public is its incomprehensibility. The technological world is a small, exclusive stage which holds only a few actors: experts. Consequently, if AI judges were implemented, only a select few (the aforementioned experts and those who have the money to pay for an expert explanation) would truly understand the way in which they work. Machine learning is the predominant technique for realizing AI. It uses techniques that lack the explicit logical reasoning or causal inferences that define human explanations. The public, save for experts, would be unable to comprehend these complex mechanisms.

This incomprehensibility has two consequences. Firstly, it erodes the legitimacy of the legal system. If the public is to accept judicial decisions, it needs to have at least a basic understanding of them. Additionally, if the process of resolving a legal dispute becomes a black box, individuals will feel disempowered and vulnerable to veiled injustice, just like Josef K in “The Trial”.

The second consequence is the reduced sense of accountability in the law. If individuals cannot understand the reasoning behind the decision, they will not be able to challenge it or spot any errors, therefore allowing potential injustice to fester. For instance, how can society detect whether the program used by an AI judge incorporates the racism of the programmer, if they are unable to understand its inner workings?

---

9 https://towardsdatascience.com/clearing-the-confusion-ai-vs-machine-learning-vs-deep-learning-differences-fce69b21d5eb
10 Benjamin N. Cardozo, “The Nature of the Judicial Process”
11 idem 9)
Having made these points, it becomes clear that artificial means of dispute resolution are not preferable to human ones. The appeal of AI adjudication is eroded by considerations of justice and the need to preserve the legitimacy of the legal system. But one question remains: should there be a division of labor between human and AI judges? For instance, should humans be required to oversee the verdicts given by robot judges in more complex cases, or even delegate simple ones?

What seems to be the perfect middle ground can end up becoming the worst of both worlds. Combining human and artificial decision-making can maintain the legitimacy of the law, but it sacrifices efficiency. For instance, imagine that human judges have to oversee the verdicts given by AI judges before they can be considered “acceptable”. This would significantly diminish the number of cases that could be resolved if robot judges were left to their own devices.

Similarly, in the first part of my essay I have argued that technology might significantly alter the values underpinning our legal system, pushing towards a more standardized form of adjudication. The acceptability of human discretion in the law is in itself mutable. Delegating even simple cases to AI judges could have major implications in the long term, by rendering discretion not only less feasible, but also less palatable.

It ought to be conceded that the legal system is riddled with tasks that are suited for robotization, such as identifying relevant precedents or the potential outcomes of a dispute, without any social costs or the risk of modifying values. But, despite its virtues, artificial dispute resolution is a fond hope, which runs the risk of reducing the legal system to a set of standardized procedures and alienating the society it should protect. Just as Lon Fuller stated, “water from a tainted spring can sometimes be purified, but only at the cost of making it something other than it was”.

word count: 1940 (including footnotes)