Acta Univ. Sapientiae, Legal Studies, **10**, 2 (2021) 161–168

DOI: 10.47745/AUSLEG.2021.10.2.01

Can Justice Be Anything Other than Human?

Emőd VERESS
PhD, University Professor
Sapientia Hungarian University of Transylvania (Cluj-Napoca, Romania)
Department of Law

Senior Research Fellow
Mádl Ferenc Institute of Comparative Law (Budapest, Hungary)
e-mail: emod.veress@sapientia.ro

Abstract. Artificial intelligence (AI) is an emerging technology in the field of justice, which, though only used in limited fashion, is likely to bring potentially disruptive changes in the future. This study examines some of the most significant characteristics of AI that make it suited, and at the same time unsuited, to application during litigation. Specifically, the problems of opacity in decision making by machine learning algorithms and the lack of a way to project flexible human morality on such systems is emphasized. The author concludes that AI will be a useful tool for preparation of litigation; however, it is less likely to replace the human judge, as adjudication presupposes understanding of human sentiment and the sense of justice, which requires some degree of bias in a positive sense.

Keywords: artificial intelligence, human factor, litigation, judge, machine learning

One of the most significant suggestions made by those seeking to predict the future is that artificial intelligence (AI) will be employed in the judiciary, resulting in the replacement of some legal professions – as seems preferable by some, mainly of lawyers – perhaps even of judges. One of the main proponents of such views is Richard Susskind, who in his writings[1] (*The End of Lawyers?* or *Tomorrow's Lawyers: An Introduction to Your Future*) explores the potential effects of artificial intelligence on legal professions.

This direction of development and its possible impact on legal services can already be seen in Romania.[2] The former president of the National Union of Romanian Bar Associations made a particularly pessimistic statement[3] about the

1 For a relevant review and critique, see Madocsai 2014.
2 Macovei 2018.
3 Florea 2017.

risks posed to the judiciary by artificial intelligence. Reflections projecting an increasingly gloomy future in this field also concern the legal community as to whether it is possible to replace the human factor (human judge) with a machine in resolving litigation.

Like so many other issues related to artificial intelligence, this seems to be a philosophical concern that deviates from reality, but in fact it explores the technological possibilities of overcoming an often quite serious problem: the lack of impartiality or judicial bias. Just as the general public has hoped for fairer decisions in football since the introduction of video refereeing (and instead the rigidity of referees' calls has increased),[4] the technological replacement of judges (as Richard Susskind argues) can be hoped to lead to cheap, fast, incorruptible, and above all impartial justice (but may in effect lead to something entirely different).

Such utopian ponderings raise two significant issues: 1. AI employed in the judiciary will not be able to use the tools of human cognition to resolve the dispute but will instead utilize opaque IT algorithms; 2. 'mathematical' justice measured out with a pharmacist's scale constituted by an algorithm in the name of so-called complete impartiality can lead to serious injustices that are incompatible with both the human sense of right and wrong and a sense of legality in individual cases.

1. The most efficient implementation of artificial intelligence now known is the so-called machine learning, including deep learning[5] algorithms. However, AI is not the same as the concept of machine learning[6] but a much broader notion. While 'intelligence' also refers to the meaning of flexible, creative machine-based problem solving, which is commonly referred to as *general* artificial intelligence (GAI) – to which science is not much closer than it was 50 years ago – machine learning is limited to an algorithm that has a large but finite number of steps, which, by performing a repetitive operation very quickly, is able to discover certain correlations in a given mass of data. This produces apparently 'intelligent' answers to narrowly defined questions such as the identity of the contents of two photographs.

The larger the amount of data machine learning is based on and the more examples available to it from the data being queried, the more precisely the algorithm 'learns' to delineate the elements of the data mass desired to be filtered out or separated from others. In order for machine learning algorithms to work as intended, the program must therefore be 'taught' first, and the essential criteria for selection must be introduced, which it must then recognize in the data subsequently provided to it during operation. The operation of machine learning

4 Spitz–Moors–Wagemans–Helsen 2018.
5 Hof 2013.
6 Marr 2016.

is similar to first graders' learning during calligraphy exercises: the teacher draws the letter to be learned on the blackboard, and then the elementary school pupil tries to imitate it to the best of his/her abilities, copying the given character repeatedly one after the other until s/he learns to draw it with sufficient certainty.

For this reason, through machine learning, a machine that is said to be artificially intelligent can only learn to recognize a pattern over time, once it has been previously 'taught' with many examples considered demonstrably correct for the given pattern by a human operator. The recognition of the pattern is practised by the machine in such a way that the correct results it has achieved are followed by some manner of positive feedback from its human operator who reviews the results' correctness in the course of operation or subsequently. This feedback may also be achieved automatically, based on human-defined criteria during operation. Such an algorithm can, for example, select histological samples containing detectable tumours,[7] based on images of the given samples, with an accuracy well beyond that of a radiologist provided that sufficient examples (photographs) of the phenomenon to be identified are provided during the training phase. Exactly how this identification works during the running of the algorithm is better understood for some algorithms and less obvious to programmers[8] in the case of others, depending on the technology used, a phenomenon known as 'opacity'. From this stem the two main obstacles to the application of algorithms commonly referred to as artificial intelligence to the field of justice.

On the one hand, the concept of justice has not been satisfactorily defined by mankind. John Rawls, perhaps the most prominent philosopher on the subject, for example, has identified[9] several different theories of fair behaviour that serve social cooperation. The aim of the judiciary is precisely to preserve this social cooperation (the so-called social contract). At present, there is no way to give a machine sufficient examples of what is generally accepted as a socially just solution to a dispute so that, if it recognizes the defining characteristics of that dispute, it would be able to resolve it on its own with a comparable degree of justice. Indeed, it is not only individuals who value and apply the concept of justice in a heterogeneous way but also the legal norms themselves. Is it fair, for example, for a thief who steals food for his/her starving children to face the same punishment as a well-heeled kleptomaniac? Obviously not. A computer trained on random judgments in cases of theft may not be able to detect this difference.

Machine learning algorithms always try to demarcate certain elements from the dataset they are trained with, on the basis of criteria that are considered self-explanatory to them (although not obvious to others) and based on some machine-identified common characteristics.

7 Hu–Tang–Wang–Zhang–Zhang–Sun 2018. 134–149.
8 Smith 2018.
9 Rawls 1994.

As many aspects of human life can now be digitized, it is conceivable that all known information about a person (genetics, health, financial and educational information, parents' and siblings' identity, political views, banking details, written and oral recordings, criminal history, etc.) may be collected and collated into a single database and would then serve to fully describe the person. From such a dataset, the computer can be taught to identify, for example, the common behavioural characteristics of individuals convicted of certain crimes or suffering from certain illnesses. Later, the machine will be able to tell whether or not given persons are prone to commit a certain crime or contract a certain disease, based solely on the person's 'digital footprint'. To render justice though – by the method of machine learning known today – a computer would need to reach deeper insights, which it cannot yet do. What it can do is to proceed to a form of social profiling: by use of a given (typically statistical) method, it can determine the likelihood that a person is prone to certain behaviours or may have even exhibited such behaviours in the past. However, this finding cannot itself lead to a 'just' judgment, only to one based on a new form of prejudice, a problem characteristic of the human factor (the judge), but the very one which we would want to eliminate in the first place through the use of computers in the administration of justice. This operation of machine learning algorithms will have a good chance of being put to practical use for the first time in today's totalitarian societies.

On the other hand, the 'cognitive processes', the internal reasoning of the machine cannot always be made explicit in order to learn the truth about what factors it took into consideration and which ones it chose to ignore when rendering a decision. IT professionals are not always able to determine why certain machine learning algorithms produce a particular result, even if the result is from the intended operation of the algorithm. This is called the problem of 'opacity' or accountability of algorithms[10] in machine decision making. If we cannot determine exactly why an algorithm has made a particular decision (or perhaps we can only determine that it did so because it found the solution as being *most likely* for the given case), we cannot verify the fairness of the decision. This problem[11] already exists[12] although it still seems[13] significant to only a few particular situations. In law, however, accountability as well as legal and factual justification for decision making are essential. Only based on the statement of reasons given by the court can we determine if a judgment is fair and legal. Only based on the reasons can the possibility of administrative and judicial remedies be ensured to the affected person who sought justice against an illegal, unfair, or unjust decision.

10 World Wide Web Foundation 2017.
11 Diakopoulos–Friedler 2016.
12 Diakopoulos 2015.
13 Kemper–Kolkman 2019.

2. Law in the broadest sense can by no means be simplified into mathematical, logical processes because the application of law is much more human-bound and subjective than what would permit algorithms to replace the human factor. Law is a creation of the state that conveys dominant values and political goals. Imagine what would have happened if in the cases regarding the restitution of the Batthyaneum Library (in Alba Iulia, Romania) or the Székely Mikó College of the Reformed Church (in Sfântu Gheorghe, Romania) these were resolved by an AI entity, based on a serious and objective consideration of the facts, and this entity would have rendered a decision for restitution, having previously ascertained that these immovables – over which property rights disputes endured – constituted the property of the Catholic Church or of the Reformed Church, and not that of the Romanian state.

Do we dare to imagine that the application of the law would be ceded by any state, taken as a community of national interest, to an AI entity, which could then act with complete and utter impartiality? Or could we accept the alternative that a national consciousness or a nationalist pattern of behaviour may emerge in the AI entity either by inadvertent 'pollution' of the data pool used to teach it or as a result of deliberate programming choices? As society becomes more complex, the need for legal regulation is ever-growing. This may lead to norms regulating AI to either generate or avoid any of these outcomes…

A perhaps less menacing but equally significant possibility is the potentially revolutionary effect AI may have on legal research and documentation, which is still not without its perils and difficulties. For example, the advent of constructs similar to the information bubbles already extant in the case of social media platforms is not at all inconceivable and would carry inherent risks to the fairness of judicial procedures. Artificial intelligence and IT in general play an increasingly important role in the service of law enforcement, and this role will continue to expand. Artificial intelligence itself is one of the most difficult new areas of norm development, as its operation and spread is inconceivable without precise legal regulation.

The perils of artificial intelligence must be contained by legal preventive measures. Without the right legal framework, according to the experts themselves who are at the forefront of its development,[14] the use of some AI tools seems to be becoming an increasingly dangerous gamble.

The reality is that the true expectation of the state from any system instituted for the application of the law is that it remains – to a certain extent – subjective, and that which is subjective is either prone to – or indeed driven by – human prejudice. This has both negative and positive aspects. If we are to perform more complex legal tasks with AI, then these typically human emotions and prejudices must at first be mastered by the AI entities employed in order for such

14 Vincent 2018.

tendencies to be brought under their 'rational' control so as to be then considered in the service of a fairer way of rendering justice. Such a process is possible: AI develops and improves by processing huge amounts of data. When data tainted with prejudice is fed in to create an artificially intelligent process, the process shall inherit[15] the prejudice reflected in the data. However, by teaching AI entities to detect and ignore such prejudice, or in the service of justice (compatible with the human sentiment of fairness) consider other elements of the case than 'just the facts', a better system of rendering justice may truly be achieved.

Obviously, it should not be ruled out, and it should even be considered reasonable, that simple legal work can be automated, and this will lead to a certain reduction in the number of lawyers. Yet, opinions predicting the disappearance of the legal professions are unfounded. There is no need to worry about the legal professions, even as the role of AI grows and the use of IT tools becomes more and more intensive in law enforcement. The 'human factor' may only be introduced into law enforcement by projecting that human factor onto whatever solution for rendering justice exists (i.e. AI cannot entirely substitute but may effectively assist the human lawyer or judge). Let us consider without undue emotion the following: would we entrust the application of criminal law to algorithms, for example? Or can we imagine an AI entity applying Islamic religious law, using Sharia norms, as it objectively approves the institution of a fixed-term marriage (or that of a marriage contract with a minor) designed to circumvent the prohibition of prostitution?

The emergence of artificial intelligence will not eliminate conflicts in society. No technology can have such an effect – the existence of conflicts is a necessary corollary of human society. If we look to the future, we should rather fear the growth of new areas where this conflict may arise.

Imagine a dispute in which the lawyers of the parties on both sides use the help of AI. Conflict and competition between artificial intelligences may then also be an emerging new issue. One of the positions, however, must ultimately prevail, otherwise the dispute will remain unresolved. When applying the law, the (human) judge considers the partial truths existing on both sides, and absolutizes one of the partial truths in order to resolve the dispute. This is why legal and moral truths can differ in many cases. The question persists if AI is able to do the same.

Can a third artificial intelligence decide the conflict between two artificial intelligences? There is a significant difference between AI entities meant to project human abilities and a human who calls upon the help of an AI entity. Artificial intelligence will play an important role in many legal processes. For example, it can significantly reduce transaction costs for contracting. However, in the case of litigation, its role is likely to be applied mostly to preparatory and advisory tasks.

15 Zou–Schiebinger 2018.

The general danger of artificial intelligence in the long run is human dementia. Excessive use of AI might endanger legal culture and knowledge and the necessary and appropriate level of human conflict resolution. AI can only be fully exploited by a lawyer who is in control of it and highly trained in his/her own right.

'Releasing' innovation, and law as the value-based management and development system of human society, which pioneered moral processes from the hands of humanity, is truly a gamble.

References

DIAKOPOULOS, N. 2015. Algorithmic Accountability. Journalistic Investigation of Computational Power Structures. *Digital Journalism* 3. https://www.tandfonline.com/doi/abs/10.1080/21670811.2014.976411?journalCode=rdij20 (accessed on: 10.01.2020).

DIAKOPOULOS, N.–FRIEDLER, S. 2016. How to Hold Algorithms Accountable. *AIT Technology Review* 17.11.2016. https://www.technologyreview.com/s/602933/how-to-hold-algorithms-accountable/ (accessed on: 10.01.2020).

FLOREA, G. 2017. Inteligența artificială s-ar putea întoarce împotriva umanității! *Juridice.ro* 16.11.2017. https://juridice.ro/essentials/1828/gheorghe-florea-inteligenta-artificiala-s-ar-putea-intoarce-impotriva-umanitatii (accessed on: 10.01.2020).

HOF, R. D. 2013. Deep Learning. *AIT Technology Review* 23.04.2013. https://www.technologyreview.com/s/513696/deep-learning/ (accessed on: 10.01.2020).

HU, Z.–TANG, J.–WANG, Z.–ZHANG, K.–ZHANG, L.–SUN, Q. 2018. Deep Learning for Image-Based Cancer Detection and Diagnosis – A Survey. *Pattern Recognition* November 2018. https://doi.org/10.1016/j.patcog.2018.05.014 (accessed on: 10.01.2020).

KEMPER, J.–KOLKMAN, D. 2019. Transparent to Whom? No Algorithmic Accountability without a Critical Audience. *Information, Communication & Society* 2019. https://www.tandfonline.com/doi/full/10.1080/1369118X.2018.1477967 (accessed on: 10.01.2020).

MACOVEI, M. 2018. Înlocuieşte inteligența artificială juristul? *Juridice.ro* 26.08.2018. https://www.juridice.ro/598482/inlocuieste-inteligenta-artificiala-juristul.html (accessed on: 10.01.2020).

MARR, B. 2016. What Is the Difference Between Artificial Intelligence and Machine Learning? *Forbes* 12.06.2016. https://www.forbes.com/sites/bernardmarr/2016/12/06/what-is-the-difference-between-artificial-intelligence-and-machine-learning/ (accessed on: 10.01.2020).

RAWLS, J. 1994. *The Main Idea of the Theory of Justice*. https://www.csus.edu/indiv/c/chalmersk/econ184sp09/johnrawls.pdf (accessed on: 10.01.2020).

SMITH, A. 2018. Franken-Algorithms: The Deadly Consequences of Unpredictable Code. *The Guardian* 29.08.2018. https://www.theguardian.com/technology/2018/aug/29/coding-algorithms-frankenalgos-program-danger (accessed on: 10.01.2020).

SPITZ, J.–MOORS, P.–WAGEMANS, J.–HELSEN, W. F. 2018. The Impact of Video Speed on the Decision-Making Process of Sports Officials. *Cognitive Research: Principles and Implications* 3. https://doi.org/10.1186/s41235-018-0105-8 (accessed on: 10.01.2020).

SUSSKIND, R. 2016. Artificial Intelligence and the Law Conference at Vanderbilt Law School. *YouTube* 06.05.2016. https://www.youtube.com/watch?v=xs0iQSyBoDE (accessed on: 10.01.2020).

VINCENT, J. 2018. Elon Musk, DeepMind Founders, and Others Sign Pledge to Not Develop Lethal AI Weapon Systems. *The Verge* 18.07.2018. https://www.theverge.com/2018/7/18/17582570/ai-weapons-pledge-elon-musk-deepmind-founders-future-of-life-institute (accessed on: 10.01.2020).

WORLD WIDE WEB FOUNDATION. 2017. *Algorithmic Accountability. Applying the Concept to Different Country Contexts*. https://webfoundation.org/docs/2017/07/WF_Algorithms.pdf (accessed on: 10.01.2020).

ZOU, J.–SCHIEBINGER, L. 2018. AI Can Be Sexist and Racist – It's Time to Make It Fair. *Nature* 18.07.2018. https://www.nature.com/articles/d41586-018-05707-8 (accessed on: 10.01.2020).