

# Is the Robot Judge Near?

*Henry Prakken*

*Faculty of Law, University of Groningen & Department of Information and Computing Sciences, Utrecht University, The Netherlands*

*August 20, 2020*

*This is the author's own unofficial translation of H. Prakken, Komt de Robotrechter eraan? Nederlands Juristenblad 2018/207, no. 4, pp. 269-274.*

## **Introduction**

Recently there has been increasing attention for the question how artificial intelligence (AI) can change the work of legal professionals. This is mainly because of the spectacular progress in AI in recent years (machine learning, natural-language processing, self-driving cars, big-data algorithms) and also the rise of a number of successful legal AI start-ups, especially in the US. Especially algorithms that predict outcomes of legal cases have drawn the attention. Many conclude from the existence of such predictive algorithms that automated legal decision-making is now within reach. In this article I argue that this conclusion is based on misunderstandings concerning the nature of predictive algorithms and a neglect of the difference between *predicting* and *taking* legal decisions. For automating judicial decision-making (as suggested for “routine cases” by mr. Frits Bakker, then president of the Dutch Council of the Judiciary, during the Day of the Judiciary on September 28, 2017) a very different kind of AI system is needed than predictive algorithms, namely, algorithms that can apply legal knowledge to justify legal decisions. In this article I shall give a brief overview of current research on this kind of AI system and discuss to which extent they can be used in practice now or in the near future. I shall conclude that it is more realistic to aim for supporting human judges with AI than for automating judicial decision making.

## **How AI could benefit legal professionals**

The recent progress in AI has several causes: the increased computing power of computers (especially when connected in networks), breakthroughs in machine learning and natural-language processing and the ever growing digital availability of data and information. Therefore, the expectation that AI will change the nature of legal work is justified<sup>1</sup>. Legal work is to a large extent information processing (case law, statutes and other regulations, scholarly literature, case files, contracts, ...), so technology for information retrieval and natural-language processing is clearly useful for legal professionals. Moreover, algorithms are nowadays very good in recognising patterns in large amounts of data, and this can be useful for both legal practice and legal academic research. For example, it can help detecting extraneous influences on judicial decisions, as in Israel, where researchers discovered that judges in their decisions on requests from parole became less lenient as lunch-time

---

<sup>1</sup> Good overviews of what is possible now or in the near future are B. Custers & F. Leeuw (2017), ‘Legal big data. Toepassingen voor de rechtspraktijk en juridisch onderzoek’, *NJB* 2017/1854, afl. 34, p. 2449-2456 (in Dutch) and L.K. Branting (2017), ‘Data-centric and logic-based models for automated legal problem solving’, *Artificial Intelligence and Law*, Vol. 25(1), p. 5-27.

approached, up to the point that just for lunch they refused almost all requests. Such information can help in sensible planning of court sessions. There are also algorithms that predict the probability of a particular outcome of a legal case, which can be useful for solicitors or their clients in deciding whether a law suit makes sense. Moreover, data analytics can be useful for judges, for example, for knowing which amounts of damages were awarded or which punishments were imposed by other judges in similar cases.

### **The hype: confusing predicting with deciding**

Although thus the expectations are partly justified, there is currently, especially in the general press, on the internet and in popular legal publications, a ‘big legal data hype’ that gives rise to the unfounded expectation that now the robot judge is also near. For example, Jaap van den Herik claimed in *Mr.* (a Dutch magazine for lawyers) of 31 October 2016 “In 2030 computers will decide legal cases”. And Folkert Jensma wrote in the Dutch newspaper *NRC* of 28 October 2017: “Big data can replace the judge”... “Artificial Intelligence can in the near future drive cars, decide legal cases and make diagnoses because it does not have access to a couple of thousands of X-rays, like human doctors, but millions. Expert systems already exist that predict with almost 80% accuracy how the European Court in Strassbourg will decide on the question whether the Convention has been violated or not.” A bit further on Jensma writes about “perfect concept decisions from the expert machine”. About the predictive algorithm referred to by Jensma the British newspaper *The Daily Telegraph* wrote on 24 October 2016: “Computer scientists ... developed an algorithm which can not only weigh up legal evidence, but also moral considerations.”

### **The ECHR algorithm**

Let us have a closer look at the prediction algorithm for decisions of the European Court of Human Rights<sup>2</sup>. Jensma calls it an expert system, but an expert system is an AI system that possesses knowledge of a particular field of expertise and that can by way of reasoning apply its knowledge to the facts of a case and thus produce a reasoned advice for a particular decision. By contrast, the system discussed by Jensma is just a statistical predictive algorithm. It has no understanding whatsoever of the law or a concrete case. What it does is a far cry from weighing up legal evidence and moral considerations, as claimed by the *Daily Telegraph*. The algorithm cannot even understand the judicial decisions it analyses: all it does is counting word frequencies in these decisions and relating them statistically to the outcomes of other cases in which these words occurred. Moreover, strictly speaking it does not even predict the outcome in new cases, since it needs a substantial part of the text of the new decision in order to predict its outcome: only if the algorithm has access to the Court’s summary of the facts of the case, its procedural history, and the arguments of the parties and to a substantial part of the Court’s own arguments, can the algorithm predict the case’s outcome.

Moreover, the algorithm cannot justify its predictions. A justification by the algorithm would look as follows: “In the text of this decision as we have written it thus far we have used these words this many times; therefore, on the basis of past decisions in which these words occurred the probability that we, when continuing to write the decision, will conclude that article X of the Convention has been violated is this high”. This, of course, has nothing to do with judicial

---

<sup>2</sup> N. Aletras, D. Tsarapatsanis, D. Preotiuc-Pietro & V. Lampos, ‘Predicting judicial decisions of the European Court of Human Rights’, *PeerJ Computer Science* 2016, 2:e93.

decision making and legally justifying the decisions. Incidentally, this shows an additional problem of predictive algorithms, namely, that all they do is producing probabilities. This is useless for a judge in a concrete case, just as the statistic that 92% of the accused in a criminal case is convicted is useless for a judge in a new criminal case. Every new case has to be assessed on its legal and factual merits.

Another question is how good is an accuracy of 79%. That is rather disappointing. To start with, the algorithm has to know about the decision-to-be-predicted that the decision was about the violation of a specific article of the UCHR and it can only answer: “yes, this article is violated” or “no, this article is not violated”. So it cannot answer the question “Will the Court decide that the ECHR is violated and, if so, which article?” Moreover, the algorithm answers a yes/no question, so even tossing a coin already scores 50%. Then 79%, although significantly more, is not dramatically much more. In addition, the Court’s HUDOC-database<sup>3</sup> reveals that in 84% of the cases of the kind investigated by Aletras et al. the Court concludes to a violation. So someone who consistently predicts that the Court will conclude to a violation does better than the algorithm. What is more, since the algorithm needs a substantial part of the text of the decision-to-be-predicted, I suspect that a human lawyer will, on the basis of the same text, score even better, maybe even 100%. After all, a human does not count words but understands them.

It should be said that not every predictive algorithm needs a part of the decision-to-be-predicted in order to make predictions. An algorithm that can predict 70% of the decisions of the US Supreme Court<sup>4</sup>, does so purely on the basis of information that is available in a database about the Supreme Court before the start of a case. But 30% incorrect for a yes/no question is still rather much. Moreover, this algorithm can also not assess a case on its legal merits, since it is based on data that are not about the merits of a case, such as personal characteristics, political preferences and the appointment history of the judges and trends in the Supreme Court’s decisions entered into the database by experts. A justification of this algorithm would look like: “most of us are republican, older than 60 and rather conservative in this kind of case, therefore we overturn the lower court’s decision” This is, of course, not a legally acceptable way to justify a decision.

### **Predicting is not the same as deciding**

We can draw two conclusions. First, predictive algorithms do not yet perform as well as many people think; and second, they are fundamentally unsuitable as a model of judicial decision making and justification: either they need a substantial part of the decision-to-be-predicted to make their prediction (and they cannot recognise legally relevant grounds in the Court’s decisions), or they base their predictions on extraneous factors instead of on the merits of a case. In both cases the predictive algorithms cannot justify their predictions in a legally acceptable way, while yet such is required by every legal decision-making procedure.

Big-data-algorithms can in many ways be useful for legal professionals, including judges, but they do not bring the robot judge any nearer. For automated decision making on the basis of the merits of a case a different kind of AI system is needed, namely the expert systems

---

<sup>3</sup> hudoc.echr.coe.int.

<sup>4</sup> D.M. Katz, M.J. Bommarito & J. Blackman (2017), ‘A general approach for predicting the behavior of the Supreme Court of the United States’. *PLoS ONE* 12(4): e0174698. <https://doi.org/10.1371/journal.pone.0174698>

(nowadays often called ‘knowledge-based systems’) referred to but misunderstood by Jensma. Such systems are AI systems that can reason with legal knowledge as a human legal professional would do and that can therefore justify their decisions or advice in a legally acceptable way.

## History of AI & law

How do realistic legal knowledge-based systems look like, and what is the state-of-the art in research on such systems? Some people think that AI & law is something of recent years but nothing could be farther from the truth. As an international research field it dates back 40 years and the Netherlands have been seriously involved since more than 30 years. Even research on legal predictive algorithms is not new. As early as in 1991 my then colleagues of the Computer/Law Institute of the Free University Amsterdam applied neural networks (a machine-learning technique) to case law on the issue whether a job offered to an unemployed person by the unemployment benefit office is ‘suitable employment’<sup>5</sup>.

Applied research on legal knowledge-based systems is based on a very simply model of legal knowledge and legal reasoning, which in fact boils down to the legal syllogism. Legal knowledge is represented in the computer as if-then rules, which are then by the computer applied to a case. This ‘rule-based’ model of legal reasoning can easily be automated with the help of the logic of if-then rules but every legal professional knows that as a model of legal reasoning it is too simplistic. For example, it leaves no room for interpreting legal rules as a human legal professional would do. If several alternative interpretations are possible, then a choice needs to be made in advance, at the time when the system is filled with legal knowledge. A computer cannot generate arguments for a choice between the alternative interpretations. Neither can a computer make reasoned decisions to not apply a rule in exceptional circumstances. The only way this can be done is by not following the computer’s advice; the computer cannot give reasons for why the rule should not be applied.

Despite these limitations, rule-based systems have still proved to be practically useful since, for example, civil servants applying social security law made many mistakes because of the vast amounts of social security regulations and their complex logical structure. By contrast, computers are ideally suited for retrieving stored information and for handling syntactic and numerical complexities. This turned out to be very useful for large-scale processing of administrative law, for example in processing social security applications or tax declarations and in immigration law, areas where ‘bulk’ processing of cases is daily practice, where the case input is largely factual and easy to verify, and where advanced legal reasoning is less important<sup>6</sup>.

---

<sup>5</sup> G.J. van Opdorp, R.F. Walker, J.A. Schrickx, C. Groendijk, and P.H. van den Berg (1991), ‘Networks at work. A connectionist approach to nondeductive legal reasoning’, in *Proceedings of the Third International Conference on Artificial Intelligence and Law*, p. 278-287, New York: ACM Press.

<sup>6</sup> See, for instance, T. van Engers, R. Gerrits, M. Boekenoogen, E. Glassée, and P. Kordelaar (2001), ‘POWER: Using UML/OCL for modeling legislation - an application report’, in *Proceedings of the Eighth International Conference on Artificial Intelligence and Law*, pp. 157-167, New York: ACM Press. J.S. Svensson (2002), ‘The use of legal expert systems in administrative decision making’, in *Electronic Government: Design, Applications and Management*, ed., A. Grönlund, p. 151-169, London etc: Idea Group Publishing.

However, for the reasoned application of regulations in a specific case such rule-based systems are unsuitable. This is left to the human users but to many this is the core of legal thinking. Basic research in AI & law<sup>7</sup> investigates whether the computer can make judgements in concrete cases as a human legal professional would do. On the basis of realistic models of legal knowledge and legal reasoning so-called argumentation systems (systems that can produce arguments for and against a decision) have been developed. Among other things ‘case-based reasoning’ is investigated in areas that have no clear classification or interpretation rules but only factors that are pro or con a decision and that can apply to specific cases in varying combinations and to varying degrees. Consider, for example, Article 9, Section 2 of the old Dutch Data protection act (now Article 6(4) of the GDPR) about the question whether further processing of personal data is compatible with the original purposes for which the data was obtained. Article 9 does not give a clear answer to this question but just lists a number of factors (such as the strength of the link with the original purpose, the nature of the data en how privacy-sensitive they are) that have to be weighed. Systems have been developed (especially in the US because of the common-law system) that can argue for decisions in a case on the basis of the similarities and differences between cases. For example, if in a past case it was decided that further use was compatible and the new case is at least as strong for the same decision (‘the data are less privacy-sensitive and the link with the original purpose is stronger’) then the systems will propose the same decision in the new case. When there are also differences with the past case (‘but the safeguards for the data subject are weaker’), the systems will give both arguments pro and con a decision. Some of these systems can in doing so even balance competing legal or moral principles or social consequences (‘the customer’s privacy is more important than the economic interests of the shop’).

Purely academically speaking this is fascinating research, in which fundamental questions about legal reasoning are addressed. But practical usability of argumentation systems is for the time being a different story. The problem is that it is very hard to represent explicit knowledge about things like balancing interests, empathy and sense of justice or social relations at a large scale in computer-understandable form. Moreover, issues of evidence and proof are also still too hard for computers, since addressing such issues requires vast amounts of commonsense knowledge about the world, and commonsense knowledge is in AI still a big obstacle for the development of practically usable knowledge-based systems, not just in the law<sup>8</sup>.

But could perhaps the recent breakthroughs in machine learning and natural-language processing be of help here? Custers<sup>9</sup> suggests that the problem has already been solved since IBM’s Watson-system won the American TV-Quiz Jeopardy in 2011 (in which the players must have knowledge of ‘trivia’ in a large number of areas), by applying (especially) machine learning and natural-language processing to public digital sources such as Wikipedia.

---

<sup>7</sup> See for overviews e.g.. H. Prakken (2015), ‘Legal reasoning: computational models’, in J.D. Wright (ed.) *International Encyclopedia of the Social and Behavioural Sciences*, 2nd edition. Oxford: Elsevier Ltd. en K.D. Ashley (2017), *Artificial Intelligence and Legal Analytics. New Tools for Law Practice in the Digital Age*. Cambridge University Press.

<sup>8</sup> E. Davis & G. Marcus (2015), ‘Commonsense reasoning and commonsense knowledge in Artificial Intelligence’. *Communications of the ACM*, Vol. 58 No. 9, p. 92-103.

<sup>9</sup> B.H.M. Custers (2017), ‘Kunnen computers het wetboek interpreteren?’ In: B. de Graaf & A. Rinnoy Kan (red.) *Hoe zwaar is licht? Meer dan 100 dringende vragen aan de wetenschap*, Amsterdam: Balans.

According to Custers computers would with Watson-like technology already be able to interpret the law at the level of human legal professionals. However, this claim is a gross exaggeration, since it is not backed by any evidence<sup>10</sup>.

Nevertheless, Watson still gives hope that one day advanced argumentation systems will be practically usable, since the required information would not have to be manually encoded any more in computer-understandable form by humans but could automatically be extracted from public digital sources by the computer. However, to realise this, still many breakthroughs in natural-language processing technology are needed. For example, currently there is much research on ‘argumentation mining’, which is the automatic recognition of elements of arguments in natural-language texts<sup>11</sup>. If this research is successful then computers can, for example, automatically recognise relevant factors in case law and apply these factors in new cases with the argumentation techniques developed in AI & law. This is fascinating research but the results are still modest. Moreover, in complex cases argumentation systems cannot do more than consider alternative ways to decide the case, since complex cases by definition have no clear unique outcome.

### **Automating decision-making in routine cases**

Having said so, mr. Frits Bakker of the Dutch Council of the Judiciary only suggested to automate decision making in routine cases, and routine cases by definition have a clear outcome and have no complex factual and interpretation issues. So they do not require advanced argumentation systems. On the other hand, a system must still have knowledge of the applicable laws. A rule-based system has such knowledge but that does not suffice for automatically deciding routine cases, since the facts of the case must first be determined and legally interpreted. There are cases in which the facts can be automatically determined, for example, when they can be retrieved from a government database or a digital case filing system. And there are cases in which the facts do not need to be legally interpreted. For example, in speeding cases determining violation of the speed limit suffices, and there are perhaps cases in which ticking a box of a simple form suffices. If both conditions are satisfied (the facts can be automatically determined and they do not need to be legally interpreted), then fully automated decision-making is possible. But in many cases these conditions will not both be fulfilled, not even in cases that for human legal professionals are routine. In this respect it is telling that many current digital case filing systems have a field for free-format text entry.

When a citizen enters his or her story in free-text format, can the legal interpretation of this story then be left to the citizen? No, since they will lack the required legal expertise. Neither can their legal councillors do so, since these would then become part of the judiciary, which is, of course, not acceptable. Can the computer automatically legally interpret the citizen’s story? This is far from trivial, even in routine cases. What is obvious for a human, is often far from obvious for a computer. Consider, for example, an online complaints system for internet fraud.<sup>12</sup> How can such a system automatically recognise that ‘cunning tricks’ were used? A

---

<sup>10</sup> See, for instance, section 1.4 of K.D. Ashley (2017), *Artificial Intelligence and Legal Analytics. New Tools for Law Practice in the Digital Age*, Cambridge University Press.

<sup>11</sup> R. Mochales Palau & M.-F. Moens (2011). ‘Argumentation mining’. *Artificial Intelligence and Law*, Vol. 19(1), p. 1-22. M. Lippi & P. Torroni (2016), ‘Argumentation mining: State of the art and emerging trends’. *ACM Transactions on Internet Technology*, Vol. 16(2), p. 10:1-10:25.

<sup>12</sup> My colleegas in Utrecht have developed such a system in collaboration with the Dutch National Police. See F.J. Bex, B. Testerink & J. Peters (2016), A.I. for online criminal complaints: from natural

rule-based system would, for doing so, need further interpretation rules that legally classify the natural-language text entered by the citizen. Even in routine cases this is extremely difficult, since natural language gives humans a nearly infinite number of ways to express themselves. And natural-language processing technology is not yet advanced enough to give correct legal interpretations of what citizens enter in ordinary language. In sum, according to the present state-of-the art in AI, fully automated decision making in routine cases is still impossible, unless the facts can be automatically determined and do not need to be legally interpreted.

### **Computer and human collaborate**

What is then possible? It is more realistic to aim for supporting human decision makers with AI, to make human and computer *together* perform better than either the human or the computer *alone*. In the short term the legal big-data-algorithms that are currently so often in the news can be useful in many ways. Some time later knowledge-based support of legal reasoning and decision-making will become increasingly possible. It has always been a dream of AI & law researchers that the computer would one day be an intelligent assistant of human legal professionals: a ‘sparring partner’ against which humans can test their thoughts about a case. Computers would, for instance, be able to suggest or compare arguments in interpretation or evidential issues and they would be able to identify weak spots in arguments. By intelligently combining the fruits of 40 years of AI & law research with machine learning and natural-language processing technology, especially in the area of argumentation mining, this dream could soon become true.

Justice-seeking citizens can also be supported with AI. For example, the online system for complaints of internet fraud could after a check with natural-language processing techniques say “in your story nothing seems to indicate the use of cunning tricks; please have a second look.” The citizen could then ask the system to give more information about the interpretation of this concept. Thus citizen and computer could together write a better complaint than citizen or computer alone.

In sum, at present much is already possible and in the future more will become possible. However, one should be cautious not to be carried away by the current hype: to make the combination of human and machine legally more smart than human or machine alone, much research is still needed, and the AI & law community is currently very active in doing this research.