Predictive Justice in the UK
(about [Viola, Interpretation of the law through mathematical models, London 21 june 2019](#))

Iain Grant **MITCHELL**
(expert in IT law, member of the CCBE IT Law Committee, member of the IT Panel of the Bar Council of England and Wales, Chair of the Scottish Society for Computers and Law, Barrister in both Scotland and England & Wales)

Legal processes can easily be reduced to algorithms. Thus, case management can be handled by algorithms, and, indeed such algorithms are now common in back office systems (whether in courts or solicitor's offices). So far, so (relatively) uninteresting.

However, we are concerned today with the much more tricky business of *predictive* justice. For a system to use a mathematical model to predict an outcome involves complex algorithms making use of machine learning, which, in turn, requires datasets to educate the system. The core question is: how far may machine learning algorithms be developed to perform tasks which are presently the exclusive domain of human actors?

Although readily susceptible to the use of case management software, procedural justice is by no means exclusively the domain of rules which are applied as a rigid formula. It commonly happens that lawyers or clients fail to meet a procedural requirement. Thus, judges are often given discretion, under the rules, to forgive such failures, with or without imposing conditions on that forgiveness. The exercise of that discretion requires more than a simple algorithm. To mimic human discretion calls for the use of machine learning algorithms, fed with a dataset containing previous exercises of discretion by human judges in similar situations. This amounts to predictive justice. The machine in effect predicts what a human judge would be expected to do, and then does likewise.

Systems which claim to be able to predict what courts can do are now commercially available and in use as a research tool in large law firms, but they remain imperfect. A recent University of London study found a degree of accuracy of only 79% in a system designed to predict the outcome of cases in the European Court of Human Rights.

We have no systems in the UK charged with actually making discretionary decisions, and that is not surprising.

First, there is the problem of public acceptance: not for nothing does the GDPR give a right to demand a review of an automatic decision by a human decision maker; second, what comes out of the system is only as good as what goes in, which is itself a function of how extensive and how comprehensive the dataset may be, and whether it contains any unconscious biases.

First, comprehensiveness: judges are making discretionary decisions every day; and most of them are of little legal interest and so are not reported. Indeed the judges may not even issue a written Opinion. The ones that see the light of public gaze are the outliers, not the everyday decisions needed to populate a dataset. Second, bias: it has been suggested that almost no selection of data is without some form of bias, even if it is unconscious. For example, a relevant factor in deciding whether to grant bail is the likelihood of the accused offending if released. In some states of the United States, judges use algorithms to assess that likelihood. It is notorious that such systems tend to be biassed against young black males from socio-economically disadvantaged backgrounds. No-one programmed the algorithms to have that bias, but the dataset used to educate the system reflects the fact that police and prosecutors tend to show bias in whom to arrest and whom to prosecute. Happily, we do not use such systems in the UK

So far I have spoken about procedural justice. In the area of substantive law, you would be correct to suppose that the same problems are present; but substantive law systems also bring greater challenges.

Substantive law is not, generally, a set of simple rules, which can be automatically applied. In codified systems, it may be that there is some clarity as to what the underlying rules may be – Sig. Viola's book shows how algorithms can be devised to solve legal problems in a system where the underlying rules are clear, and where there is a clear hierarchy of rules to enable the interpretation of the law, but it is difficult to apply those algorithms to common law systems where it can often be a matter of dispute what the law

actually is. For example, judges are still arguing about whether there is an implied underlying obligation of good faith in English contract law.

The challenges of creating predictive algorithms in such a climate are immense. A predictive justice algorithm will tend to aim for the central part of the distribution curve of probable judicial outcomes. This, in effect, reinforces received wisdom. It is not what the law, or even the law applied to the facts may be, but the average of what judges believe the law is. A danger of this approach is that the law itself becomes ossified. This a particularly acute danger in a common law system where, more than in a codified system, the law can be developed through radical innovation by advocates and judges. Innovation comes through breaking free of the old judicial consensus - a snail in a ginger beer bottle leading to a jurisprudence of liability founded on the principle of a duty to one's neighbours (*Donoghue v Stevenson [1932] AC 562)*; and, a neighbourhood dispute leading to the declaration of a new servitude of parking (*Moncrieff v Jamieson 2008 SC (HL) 1*).

We are a very long way from developing so-called "strong AI," Such Artificial Intelligence systems as we have, do nit use human deductive processes. They tend to be black boxes making use of algorithms merely to mimic the outcomes of human intelligence, but they are not the reality of human intelligence. They are not self-aware nor guided by a moral compass. As Artificial Intelligence is (to use Beaudrillard's terminology) a simulacrum, so, too, is predictive justice.

As we look at the future with its challenges and opportunities, it is as well that we have a clear-eyed appreciation of the risks, on the one hand, of being unwilling to use AI where it can reasonably be used, and, on the other, of surrendering to the *simulacra* by thinking that we can leave it all to the algorithms.